

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: December 1, 2004, 22:49:23 ; Search time 107 Seconds
(without alignments)
2238.653 Million cell updates/sec

Title: US-09-867-570-2
Perfect score: 1763
Sequence: 1 MESKSSWVIRLGLSLMDST.....EGGGMLEPQETLELSGRLEQ 337

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xidp
-Q=/cgn2_1/USPTO.spool.P/US09867570/runat.01122004.161123.6373/app.query.fasta_1.519
-DB=Issued Patents NA -QPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09867570 @CGN 1.1 105 @runat.01122004.161123.6373 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1642	93.1	969	4	US-09-254-227A-4 Sequence 4, Appli
2	1598	90.6	969	4	US-09-254-227A-6 Sequence 6, Appli
3	1405	79.7	969	4	US-09-254-227A-8 Sequence 8, Appli
4	1382	78.4	969	4	US-09-254-227A-10 Sequence 10, Appli
5	1382	78.4	969	4	US-09-254-227A-12 Sequence 12, Appli
6	1360	77.1	969	4	US-09-254-227A-14 Sequence 14, Appli
7	834.5	47.3	1011	4	US-09-254-227A-2 Sequence 2, Appli
8	529	30.0	291	4	US-09-495-050A-164 Sequence 164, App
9	466	28.1	275	4	US-09-016-434-330 Sequence 330, App
10	461	26.1	2416	4	US-09-016-434-1264 Sequence 1264, Ap
11	430.5	24.4	1388	4	US-09-016-434-1225 Sequence 1225, Ap
12	430.5	24.4	1388	5	PCT-US93-06251-26 Sequence 26, Appli

13	392.5	22.3	1327	6	US-09-464-970B-134 Patent No. 5320941
14	381.5	21.6	2435	4	US-08-981-825-5 Sequence 134, App
15	255.5	14.5	2911	3	US-09-480-784-5 Sequence 5, Appli
16	255.5	14.5	2911	3	US-09-480-784-5 Sequence 5, Appli
17	240	13.6	1200	5	PCT-US95-03032-1 Sequence 1, Appli
18	238	13.5	1062	4	US-09-016-434-1207 Sequence 1207, Ap
19	231	13.1	1200	5	PCT-US92-02977-1 Sequence 1, Appli
20	227.5	12.9	1106	5	PCT-US92-02977-5 Sequence 5, Appli
21	227.5	12.9	1106	5	PCT-US95-03032-4 Sequence 4, Appli
22	226.5	12.8	1510	4	US-07-759-568-4 Sequence 1140, Ap
23	226.5	12.8	1510	4	US-09-023-655-1140 Sequence 8, Appli
24	226.5	12.8	1748	4	US-08-202-056-8 Sequence 1135, Ap
25	226.5	12.8	1750	4	US-09-023-655-1135 Sequence 1, Appli
26	225.5	12.8	1143	1	US-08-467-125-1 Sequence 1, Appli
27	225.5	12.8	1143	3	US-08-911-320A-1 Sequence 1, Appli
28	225.5	12.8	1143	3	US-09-217-101-1 Sequence 1, Appli
29	225.5	12.8	1842	4	US-09-016-434-46 Sequence 46, Appli
30	220	12.5	1244	1	US-07-816-283-7 Sequence 7, Appli
31	220	12.5	1244	1	US-08-417-103-7 Sequence 7, Appli
32	220	12.5	1373	5	PCT-US92-02977-6 Sequence 6, Appli
33	220	12.5	1373	5	PCT-US95-03032-3 Sequence 3, Appli
34	217	12.3	1058	4	US-09-016-434-1206 Sequence 1206, Ap
35	217	12.3	1910	4	US-09-944-807-1 Sequence 1, Appli
36	217	12.3	2631	4	US-09-023-655-1083 Sequence 1083, Ap
37	214.5	12.2	1317	4	US-09-016-434-1446 Sequence 1446, Ap
38	214	12.1	1438	4	US-09-016-434-1426 Sequence 1426, Ap
39	211	12.0	1062	4	US-09-576-1608-12 Sequence 12, Appli
40	210	11.9	1068	4	US-09-170-496D-1 Sequence 1, Appli
41	206	11.7	1068	4	US-09-170-496D-163 Sequence 163, App
42	205.5	11.6	1289	4	US-09-745-842-1 Sequence 1, Appli
43	205	11.6	1002	4	US-09-170-496D-7 Sequence 7, Appli
44	205	11.6	1265	4	US-09-016-434-1430 Sequence 1430, Ap
45	204	11.6	1002	4	US-09-170-496D-167 Sequence 167, App

ALIGNMENTS

★
RESULT 1
US-09-254-227A-4
Sequence 4, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Paola
APPLICANT: Shi-Hsiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 969
TYPE: DNA
ORGANISM: Homo sapiens
US-09-254-227A-4
Alignment Scores:
Pred. No.: 1.07e-157
Score: 1642.00
Percent Similarity: 98.76%
Best local Similarity: 97.20%
Query Match: 93.14%
DB: 4
Gaps: 0
US-09-867-570-2 (1-337) x US-09-254-227A-4 (1-969)
Cy 16 MetaspserthrileprovalleuglythrGlueuthrProfilaeanglyarxGluglu 35
Db 1 ATGATCAACACATCCAGTCTTGAGTACAAACACACACATCAACGAGCGTGGAGAG 60

QY	36	ThrProCysTylLeuSglnThrLeuSerPheThG1VleuThrCysIleValSerLeuVal	55
Db	61	ACTCTTGGTACACCAACCCTGACCTTACCGGGCTGACGTGCATCTATTTCCCTTGTG	120
QY	56	AlaLeuThrG1YAsnAlaValValLeuTrpLeuLeuG1CYsArgMetArgArgAsnAla	75
Db	121	GCGGTGACAGAAACGCGGGTGTGCTGTGGCTCCTGGGCTGCGGACATGCCAGAAAGCT	180
QY	76	ValSerIleTyrI1leLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerG1YHisIle	95
Db	181	GTCTCCATCTTACATCTTCAACCTGATCGGGCCAACTCTCTTCTTCTTACCGGCACATT	240
QY	96	IleCysSerProLeuArgLeuIleAsnI1eArgH1sProI1eSerI1eLeuSerPro	115
Db	241	ATATTTTGGCCGCTTACCCCTCATCAATATTCGCCATTCATCTCCAAATATCTCAAGTCT	300
QY	116	ValMetThrPheProTyrPheI1eG1YLeuSerMetLeuSerAlaI1eSerThrGluArg	135
Db	301	GTGATGACCTTTCCCTACTTTATAGGCGCTTAGCATGTGACGGCGCATCAGACCGAGCG	360
QY	136	CysLeuSerI1eLeuTrpProI1eTrpTyrHisCysArgArgProArgI1YLeuSerSer	155
Db	361	TGCCCTGTCATCTCTGGCGCCATCTGTATGCACATGCGCGCCGCCAGATACCTGTCAATCG	420
QY	156	ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerI1leLeuGluTrpMetPhe	175
Db	421	GTCATGTGTCTCTGCTGTGGGCCCTGTCTCTGCGAGATTCCTGGAGTGGATTTTC	480
QY	176	CysAspPheLeuPheSerG1YAlaAspSerValTrpCysGluThrSerAspPheI1eThr	195
Db	481	TGTGACTTCTCTGTTTAGTGGTGAATTCGTTTGGTGGTGAAGCATTCATTACACA	540
QY	196	I1eAlaTrpLeuValPheLeuCysValValLeuCysG1YSerSerLeuValLeuLeuVal	215
Db	541	ATCGGGTGGCTGATTTTTTATATGTGTGTTCTCTGTGGGTCCAGCCGTGCTGTGGTGC	600
QY	216	ArgI1eLeuCysG1YSerArgI1YMetProLeuThrArgLeuTyrValThri1eLeuLeu	235
Db	601	AGGATTTCTGTGGATCCCGGAGATGCCCTACACAGGCTGTATCGTACCATCTCTTC	660
QY	236	ThrValLeuValPheLeuLeuLeuCysG1YLeuProPheG1YI1eGlnTrpAlaLeuPheSer	255
Db	661	ACAGTGTGGTCTTCTCTCTCTGTGGCGTCCCTTTGGCATTCAGTGGGCGCTGTTTCC	720
QY	256	ArgI1eHisLeuAspTrpI1YsValLeuPheCysHisValHisLeuValSerI1ePheLeu	275
Db	721	AGGATCCACCTGGATTTGGAAGCTTATTTTGTCACTGTGATCTAGTTTCCATTTTCTCG	780
QY	276	SerAlaLeuAsnSerSerAlaAsnProI1eI1eTyrPhePheValG1YSerPheArgGln	295
Db	781	TTCGGCTTTACAGCAGTGCACCAACCCCATATTACTCTTCTGTGGGCTCTCTTAGGCAG	840
QY	296	ArgGlnAsnArgGlnAsnLeuYsLeuValLeuGlnArgAlaLeuGlnAsnProGln	315
Db	841	CGTCAAAATAGGCAAAACCTGAAGCTGTCTCCAAAGGCTCTGCAGACAGCGCTGAG	900
QY	316	ValAspGluG1YG1YTrpLeuProGlnGluThrLeuGluLeuSerG1YSerArgLeu	335
Db	901	GTGTGTGAAGGTGGAGGTGTGCTTCTCCAGGAACCTCGAGCTGTGTGGAGAACAAATTG	960
QY	336	GluGln	337
Db	961	GAGCAG	966

```

/ APPLICANT: Lembo, Paola
/ APPLICANT: O'Donnell, Dejan
/ APPLICANT: Shi-Hsiang, Shen
/ TITLE OR INVENTION: G Protein-coupled Receptors from the Rat and Human
/ FILE REFERENCES: 81823/268117
/ CURRENT APPLICATION NUMBER: US/09/254,227A
/ CURRENT FILING DATE: 1999-03-03
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6
/ LENGTH: 969
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-254-227A-6

Alignment Scores:
Pred. No.: 5,5e-153 Length: 969
Score: 1598.00 Matches: 307
Percent Similarity: 97.20% Conservative: 5
Best Local Similarity: 95.64% Mismatches: 9
Query Match: 90.64% Indels: 0
DB: 4 Gaps: 0

```

US-09-867-570-2 (1-337) X US-09-254-227A-6 (1-969)

QY	1	MeAspSerThrIleProValLeuGIYhngIleuLeuTrpIleAsnGIYArgGIuGI	35
Db	1	ATGATATCCAAACGCTCCCAAGTCTTGGGTACAGAACTGACCAATCAACGACGTGAGAG	60
QY	36	ThrProCysTyrIleGlnThrIleuSerPheThrGlyLeuThrCysIleValSerLeuVal	55
Db	61	ACTCCCTGGCTACAGAGAGACCTGAGCTTCAACGGGGCTGAGCGTCAATCGTTTCCCTTGT	120
QY	56	AlaLeuThrGlyAsnAlaValValLeuLeuTrpLeuLeuGIYcysArgMetArgArgAsnAla	75
Db	121	GGCCTACAGGAACCGGTTGTCTCTGGCTCTCGGGCTCCGCAATCGCAGAAAGCT	180
QY	76	ValSerIleTyrIleLeuAsnLeuValAlaAlaPheLeuPheLeuSerGlyValIle	95
Db	181	GTCCTCATTCATCTCTCAACCTGGATCCGGCCGCACTTCTCTTCTTACGGGCACATTT	240
QY	96	IleCysSerProLeuArgIleuLeuAsnIleAsnIleAsnProIleSerArgIleLeuSerPro	115
Db	241	ATATGTTGCCCGTTAGCGCTCATCAATACAGCAATCCCATCTCCAAATCTTAGTACT	300
QY	116	ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg	135
Db	301	GTGATGACCTTCCCTACTTTATAGCGCTTAAGCAGTGTGAACGGCATCAGCAGCAAGCGC	360
QY	136	CysLeuSerIleLeuTrpProIleTrpTyrIleCysAlaGlyGlyProAlaGlyTyrLeuSerSer	155
Db	361	TGGCTTTCATCTGTGGGCCATCTGGTACACATCGCCGCCCCCAATACCTGTCAATCG	420
QY	156	ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe	175
Db	421	GTCATGTGTCTCTGTCTCGGGCCCCGTCCTCGCTCGCGAATATCTCGAGTGTGATTTTC	480
QY	176	CysAspPheLeuPheSerGlyValAlaPheSerValTyrCysGluThrSerAspPheIleThr	195
Db	481	TGTGACTTCTCTGTTTATGTGGTCTCATATTCTGTTCGGTGTGAAGCGTACAGATTTCATTACA	540
QY	196	IleAlaThrPheValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal	215
Db	541	ATGGCTGGCTGGTTTTTTTACGTGTGATTTCTCTGTGGGTCCAGCTCGTCTCTGTCTC	600
QY	216	ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu	235
Db	601	AGGATTTCTCTGTGATATCCCGAAGATGCGCGCTGACAGAGGTGTACGTACATCTCTCTC	660
QY	236	ThrValLeuValPheLeuLeuCysGlyLeuLeuProPheGlyIleGlnTrpAlaLeuPheSer	255
Db	661	ACAGTGTGTCTCTCTCTCTGTGTGCTCGCCCTTTGGCATTCAGTGGGCCCTGTATTTC	720

RESULT 2
US-09-254-227A-6
Sequence 6, Application US/09254227M
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmed, Sulcan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves

Qy	1	MeAspSerThrIleProValLeuGIYhNrgIleuThrProIleAsnGIYArgGIuGIu	35
Db	1	ATGATATCCAAACGCTCCCAAGTCTTGGGGTACAAATGACATGACCAATTCMAAGACGTGAGAG	60
Qy	36	ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal	55
Db	61	ACTCCTTGGCTACAAAGACACCTGAGCTTCAACGGGGCTGAGGTGATCATGTTCCCTTGTGC	120
Qy	56	AlaLeuThrGlyAsnAlaValValLeuThrPheLeuGluYcAspGmetArgAsnAla	75
Db	121	GGCGTACAGAAACCGGTGTGTGCTCTGGCTCCCTGGGCTCCGCAATGGAGAAAGCT	180
Qy	76	ValSerIleTyrIleLeuAsnLeuValAlaIlaPheLeuPheLeuSerGlyHleIle	95
Db	181	GTTCCATATCAATCTCTCAACCTGGATCGGGCGACATTCTCTTCTTTAGGGCCACATT	240
Qy	96	IleCysSerProLeuArgLeuIleAsnIleArgHleProIleSerGlyIleLeuSerPro	115
Db	241	ATATGTTCGCCGTATCGCTCATCAATATCAAGCCATCCCATCTCCAAATCTTCAGTCCCT	300
Qy	116	ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg	135
Db	301	GTATATACCTTTCCTACTTATTAAGCCTAAGACATGCTGAACGGCATATGACACCGACGCG	360
Qy	136	CysLeuSerIleLeuThrProIleTrrPyHleCysArgArgProArgTyrLeuSerSer	155
Db	361	TGCTGTTCATCTGTGGGCGGCATCTGGTACACATCGCGCGCCGACAGTACTGTATCG	420
Qy	156	ValMetCysValLeuLeuThrPAlaLeuSerLeuLeuArgSerIleLeuGluTrrPhePhe	175
Db	421	GTCATGTGTGCTCTGTCTGGGCCCGCCGTCCTGTGCGAGATATCTCGAGTGGATGTTTC	480
Qy	176	CysAspPheLeuPheSerGlyAlaAspSerValTrrPyGluTrrSerAspPheIleThr	195
Db	481	TGTGACTTCTCTGTTTAGTGTCTTAATTTCTGTCGGATGTGAAGCTGAGATTTTCAATACA	540
Qy	196	IleAlaTrrPheValPheLeuCysValValLeuCysGlyIleSerSerLeuValLeuLeuVal	215
Db	541	ATCGCTGGCTGGTTTTTTTACGTGTGGTTCCTGTGGGTCCACAGCTGTGCTGTGCTGC	600
Qy	216	ArgIleLeuCysGlyIleSerArgIleMetProLeuThrArgLeuTrrValTrrIleLeuLeu	235
Db	601	AGGATTTCTGTGGATCCCGAAGATGCGGTGACACAGCGTACGAGACCATCTCTCTC	660
Qy	236	ThrValLeuValPheLeuLeuCysGlyLeuPrrPheGlyIleGlnTrrPAlaLeuPheSer	255
Db	661	ACAGTCTGTCTTCTCTCTCTGTGTGCTGCGCCTTTGGCATTCAGTGGGCCCTGTTTTCC	720

Qy 256 Arg1LeuHisLeuAptRpyLysValLeuPheCysHisValHisLeuValSer1LePheLeu 275
 Db 721 AGGATTCACCTGGATTCGAAAGCTTTATTTGTCATGTCAGATTCAGATTTTCCTG 780
 Qy 276 SerAlaLeuAnSerSerAlaAnPro1Le1eTyRPhPheValGlySerPheArgLn 295
 Db 781 TCCGCTTTAAACAGCATGTCACACCCCATCTTACTTCTTCAGAGGCTCTTTAGGCG 840
 Qy 296 ArgGlnAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsnProGlu 315
 Db 841 CTTCAAAGACGAAAGCCCTCAAGCTGATCTCCAGAGGATTCGACAGACACCGCTGAG 900
 Qy 316 ValAspGluGlyGlyTTPLeuProGlnGlnThrLeuGlnLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAAGTGATGATGCTGCTTCACGAAACCTCGAGCTGCGGAGAACAAATTG 960

RESULT 3 US-09-254-227A-8 ; Sequence 8, Application US/09254227A ; Patent No. 6696257 ; GENERAL INFORMATION: ; APPLICANT: Ahmad, Sultan ; APPLICANT: Banville, Denis ; APPLICANT: Fortin, Yves ; APPLICANT: Lembo, Paola ; APPLICANT: O'Donnell, Dajan ; APPLICANT: Shi-Heiang, Shen ; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human ; FILE REFERENCE: 81823/268117 ; CURRENT APPLICATION NUMBER: US/09/254,227A ; CURRENT FILING DATE: 1999-03-03 ; NUMBER OF SEQ ID NOS: 22 ; SOFTWARE: PatentIn version 3.0 ; SEQ ID NO 8 ; LENGTH: 969 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-254-227A-8

Alignment Scores:
 Pred. No.: 2.18e-133 Length: 969
 Score: 1405.00 Matches: 270
 Percent Similarity: 89.44% Conservative: 18
 Best Local Similarity: 83.85% Mismatches: 34
 Query Match: 79.69% Indels: 0
 DB: Gaps: 0

US-09-867-570-2 (1-337) x US-09-254-227A-8 (1-969)

Qy 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 Db 1 ATGGATTCACCGCTCAACCTTGACACAGAACTGACCAACATCAACGAACTGAGAG 60
 Qy 36 ThrProCysTyRArgGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTTTGGTACAGACGACCTTGAGCTCAGCGTGTGACGTCGATCGTTTCCCTTGC 120
 Qy 56 AlaLeuThrGlyAsnAlaValLeuThrPheLeuGlyCysArgMetArgAsnAla 75
 Db 121 GGGCTGACAGAAACGACGTTGTACTGCTGCTGCGCTGCCATGCGCAGGAGAGCC 180
 Qy 76 ValSer1LeTyR1LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHis1Le 95
 Db 181 TTCTCATCTACATCTCAACTTGGCCGACAGAACTTCTTCTTCTCAGGCGCGCTT 240
 Qy 96 IleCysSerProLeuArgLeuIleAsn1LeArgHisProIleSer1LeLeuSerPro 115
 Db 241 ATATATTCCTCGTTAAGCTTCATCAGTATCCCATGACATCCATTAATAATCTTATCCT 300

Qy 116 ValMetThrPheProTyRPhLeuGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGATGTTTTCCTTACTTTCAGAGCTGAACTTTCGTGATGCGGAGACACGATCGC 360
 Qy 136 CysLeuSer1LeLeuThrProIleTTPTyRHisCysArgRgpProArgTyRLeuSerSer 155
 Db 361 TCCCTGCTCCCTCTGTGGCCCATCTGTACCGCTGACACGCCCCACACACTCTACCG 420
 Qy 156 ValMetCysValLeuLeuThrPalAlaLeuSerLeuLeuArgSer1LeLeuGluThrMetPhe 175
 Db 421 GTGATGATGATCTGCTGCTGAGCCCTGCTCCCTGCTGGAGACATCTCGAATGATGATTA 480
 Qy 176 CysAspPheLeuPheSerGlyAlaAspSerValTTPCysGluThrSerAspPhe1LeThr 195
 Db 481 TGTGGCTTCCTGTCACATGATGATGCTGATTCGTTGGTGTCAAACATGATTCATACACA 540
 Qy 196 IleAlaThrLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal 215
 Db 541 GTCCGCTGCGTGAATTTTATATGTGTGTTCTCTGTGGTCCAGCTGCTGCTGCTGATC 600
 Qy 216 Arg1LeuLeuCysGlySerArgLysMetProLeuThrArgLeuTyRValThr1LeLeuLeu 235
 Db 601 AGGATTCCTGTGGATCCCGAAGATACCGCTGACAGGCTGTGACGATCATCTGCTC 660
 Qy 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGly1LeGlnThrPalAlaLeuPheSer 255
 Db 661 ACAGTACTGATCTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Qy 256 Arg1LeuHisLeuAptRpyLysValLeuPheCysHisValHisLeuValSer1LePheLeu 275
 Db 721 TGATTCACCTGGACAGGAAAGCTTTATTTGTCATGTCAGATTCAGATTTTCCTG 780
 Qy 276 SerAlaLeuAnSerSerAlaAnPro1Le1eTyRPhPheValGlySerPheArgLn 295
 Db 781 TCCGCTTTAAACAGCATGTCACACCCCATCTTACTTCTTCAGAGGCTCTTTAGGCG 840
 Qy 296 ArgGlnAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsnProGlu 315
 Db 841 CTTCAAAATGAGCAAGAACCTGAACTGCTTCTCCAGAGGCTCTGACAGACACCGCTGAG 900
 Qy 316 ValAspGluGlyGlyTTPLeuProGlnGlnThrLeuGlnLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAAGTGATGATGCTGCTTCACGAAACCTCGAGCTGCGGAGAACAAATTG 960

RESULT 4 US-09-254-227A-10 ; Sequence 10, Application US/09254227A ; Patent No. 6696257 ; GENERAL INFORMATION: ; APPLICANT: Ahmad, Sultan ; APPLICANT: Banville, Denis ; APPLICANT: Fortin, Yves ; APPLICANT: Lembo, Paola ; APPLICANT: O'Donnell, Dajan ; APPLICANT: Shi-Heiang, Shen ; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human ; FILE REFERENCE: 81823/268117 ; CURRENT APPLICATION NUMBER: US/09/254,227A ; CURRENT FILING DATE: 1999-03-03 ; NUMBER OF SEQ ID NOS: 22 ; SOFTWARE: PatentIn version 3.0 ; SEQ ID NO 10 ; LENGTH: 969 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-254-227A-10

Alignment Scores:
 Pred. No.: 4.73e-131 Length: 969

Score: 1382.00 Matches: 268
 Percent Similarity: 88.51% Conservative: 17
 Best Local Similarity: 83.23% Mismatches: 37
 Query Match: 78.39% Indels: 0
 DB: 4 Gaps: 0

US-09-867-570-2 (1-337) x US-09-254-227A-10 (1-969)

QY 16 MetAapSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
 Db 1 ATGATTCACACCGCTCTCAACCTTGAACACAGAAATTACACCAATCAACCGAACTGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTTGTGACAGACAGACCTTGAAGCTCAGAGCTGACAGGCTGACGCTTCTTCTGTC 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuThrPleuLeuGlyCysArgMetArgAsnAla 75
 Db 121 GGGCTGACAGAAACCGGCTTGTGCTCTGGCTCTGGCTGCGGAGTGCAGAGAACGCC 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 TTCTTCATCTACATCATCTCAACTGGCCGACGAGACTTCTCTCTCCAGCGCCGCTT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATATATTCCTGTTAAGCTTCATCAGATACCCCAATACCAATCTTAAATCTCTATCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGATGTTTCTCACTTTCAGAGCTGAGCTTCTGAGTCCGCTGAGACCGAGCGC 360
 QY 136 CysLeuSerIleLeuThrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCGTGCTGCTCTGAGCCCATCTGGTACCGCTGCCACCGCCCAACACCTGTGAGG 420
 QY 156 ValMetCysValIleuLeuThrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
 Db 421 GTGGTGTGTCTGCTCTGAGGCCCTGCTCTGCTGCTGAGACATCTTGAGTGGATGTTA 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 Db 481 TGTGCTTCTCTGTGATGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 196 IleAlaThrPleuValPheLeuCysValIleuLeuCysGlySerSerIleValLeuLeuVal 215
 Db 541 GTCCGTGCTGATATTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
 Db 601 AGGATTCCTGTGGATCCCGGAAGATACCGCTGACAGGCTGACGTCACATCTGCTC 660
 QY 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrPalaLeuPheSer 255
 Db 661 ACAGTACGTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleHisLeuAspTyrPheValIleuPheCysHisValHisLeuValSerIlePheLeu 275
 Db 721 TGGATTCACGTGACAGGAGAGCTTATTTGTGACATTCATCTTCTTCTTCTTCTTCT 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValIleSerPheArgGln 295
 Db 781 TCCGCTCTTACAGACAGTCCCAACCCCAATCATTTCTTCTGAGGCTCTTCTTCTTCT 840
 QY 296 ArgGlnAsnArgGlnAsnLeuLeuValIleuGlnArgAlaLeuGlnAspThrProGlu 315
 Db 841 CGTCAAAATAGGACAACTGAAAGCTGTTCTCCAGAGGCTCTGACAGACGCGCTGAG 900
 QY 316 ValAspGluGlyGlyIleTyrPheProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGATGTAAGGTGAGGAGGAGCTTCTGAGGAATCTGAGCTGTGCGGAACAGATTTG 960
 QY 336 GluGln 337

Db 961 GAGCAG 966
 RESULT 5
 US-09-254-227A-12
 ; Sequence 12, Application US/09254227A
 ; Patent No. 6696257
 ; GENERAL INFORMATION:
 ; APPLICANT: Ahmad, Sulcan
 ; APPLICANT: Banville, Denis
 ; APPLICANT: Fortin, Yves
 ; APPLICANT: Lembo, Paola
 ; APPLICANT: O'Donnell, Dajan
 ; APPLICANT: Shi-Hsiang, Shen
 ; TITLE OF INVENTION: G Protein-coupled Receptors from the Rat and Human
 ; FILE REFERENCE: 81823/26817
 ; CURRENT APPLICATION NUMBER: US/09/254,227A
 ; CURRENT FILING DATE: 1999-03-03
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 12
 ; LENGTH: 969
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-254-227A-12
 Alignment Scores:
 Pred. No.: 4,73e-131
 Score: 1382.00 Length: 969
 Percent Similarity: 90.00% Matches: 268
 Best Local Similarity: 83.75% Mismatches: 32
 Query Match: 78.39% Indels: 0
 DB: 4 Gaps: 0
 US-09-867-570-2 (1-337) x US-09-254-227A-12 (1-969)
 QY 16 MetAapSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
 Db 1 ATGATTCACACCGCTCTCAACCTTGAACACAGAAATTACACCAATCAACCGAACTGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTTGTGACAGACAGACCTTGAAGCTCAGAGCTGACAGGCTGACGCTTCTTCTGTC 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuThrPleuLeuGlyCysArgMetArgAsnAla 75
 Db 121 GAGTGAACAGAAACCGGCTTGTGCTCTGGCTCTGGCTGCGGAGTGCAGAGAACGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCATCTACATCTTCAACCTTGGCCGACGACAGCTTCTTCTTCTGACGTTCCAAAT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATACGTTCCGCAATTAAGCTTCATCAATACACCAATCAATCCGCAAAATCCTGTTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGATGTTTCTCACTTTCAGAGCTGAGCTTCTGAGTCCGCTGAGACCGAGCGC 360
 QY 136 CysLeuSerIleLeuThrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCTGTCTGCTCTGAGCCCATCTGGTACCGCTGCGGCTGCGGACCAACACCTGTGAGC 420
 QY 156 ValMetCysValIleuLeuThrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
 Db 421 GTGGTGTGTCTGCTCTGAGGCCCTGCTCTGCTGCTGAGACATCTTGAGTGGATGTTA 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 Db 481 TGTGATTCCTGTTTATGATGCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCT 540
 QY 196 IleAlaThrPleuValPheLeuCysValIleuLeuCysGlySerSerIleValLeuLeuVal 215


```

Db      541  GTGCGGCGCTGATTTTATATGATGATCTCTGATTTTCCAGCCCTGCTGCTGCTC
Qy      216  Arg11eUeuCysGlySerArgLysMetProLeuThrArgLeuTyVal1Thr1leUeu 235
Db      601  AGATCTCTGTGTGGATCCCGAAGATGCCGCTGACCAAGGCTGATGTGACCATCTCTC 660
Qy      236  ThrVal1euVal1Phe1euLeuCysGlyLeuProPheGly11egIntPala1euPheSer 255
Db      661  ACAGTGTGCTCTCTCTCTCTGCGGCTCTCTCTCGGCATTTCCGAGGCGCTTAATTAC 720
Qy      256  Arg11eHis1euAspTrpLysVal1euPheCysHisVal1His1euValSer11ePheLeu 275
Db      721  AGAGTGCACCTGATTTTGAAGTCTTAATGTATGATATGATGATGATGATGATGATGATG 780
Qy      276  Ser1Ala1euAsnSerSer1AlaAsnPro11e1eTyPhePheVal1GlySerPheArg1n 295
Db      781  TCTCTCTTAACAGTATGTCACCAACCCCATCATTTACTTCTTCTGAGGCTCTTTAGGCGAG 840
Qy      296  ArgGlnAsnArgGlnAsnLeuLysLeuVal1euGlnArgAla1euGlnAspThrProGlu 315
Db      841  CGTCAAAATAGCGAAGAACTGAAAGCTGTTCTCCAGAGGCGCTGCGAGCAAGCTGAG 900
Qy      316  ValAspGlnGlyGlyTyTrpLeuProGlnGluThrLeuGlnLeuSerGlySerArgLeu 335
Db      901  GTGGATTAAGGTGAAGGCGAGCTTCTCGAGAAAGCCTGAGAGCTGTGCGAAGGAGATTG 960

```

RESULT 6

```

US-09-254-227A-14
; Sequence 14, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-254-227A-14

```

Alignment Scores:

```

Pred. No.:      8,11e-129      Length:      969
Score:          1360.00      Matches:      264
Percent Similarity: 89.06%      Conservative: 21
Best Local Similarity: 82.50%      Mismatches: 35
Query Match:    77.14%      Indels:      0
DB:              Gaps:      0

```

US-09-867-570-2 (1-337) x US-09-254-227A-14 (1-969)

```

Qy      16  MetAspSerThr1leProVal1leuGly1ThrGluLeuThrPro11eAsnGlyArgGluGlu 35
Db      1  ATGATTCACACCGTCCAGTCTTCGGTACAAACTGACACCAATTCACACGACGAGAGAG 60
Qy      36  ThrProCysTyTrpGlnThrLeuSerPheThrGlyLeuThrCysVal1SerLeuVal 55
Db      61  ACTCTCTGCTACATGACACCTGAGCTTCAAGGCTGTGACGTCGATCATTTCTCTTGTG 120
Qy      56  AlaLeuThrGlyAsnAlaVal1leuTrpLeuGlnGlyCysArgMetArgAsnAla 75
Db      121  GGAAGTACAGAGAAACGGGTTGTGCTCTGCTCTGAGGCTACCGCATGCCAGGAAGCT 180
Qy      76  ValSer1leTyTr1leLeuAsnLeuVal1Ala1AlaAspPheLeuPheLeuSerGlyHis1le 95

```

```

Db      181  GTCTCCATCTACATCTTCAACCTGGCGGAGAGACTTCTCTTCCAGCTTCCAAATT 240
Qy      96  11CysSerProLeuArgLeu1leAsn1leArgHisPro11eSerLys1leLeuSerPro 115
Db      241  ATACGTTGGCATTTAGCCCTCATCAATATACAGCATCTCATCCGCAAAATCTGTCTT 300
Qy      116  ValMetThrPheProTyTrpPhe11eGlyLeuSerMetLeuSerAla11eSerThiArg 135
Db      301  GTGATGACCTTCCCTCATCTTTACAGGCGCTGATGATGTGAGCCGCATCAGACACGAGCG 360
Qy      136  CysLeuSer11eLeuTrpPro11eTyTrpHisCysArgArgProArgTyTrpLeuSerSer 155
Db      361  TGCCTGTCTGTGTGCGCCATCTGTATCGGTACCGCTGCGCGGCCACACACTGTCAAGCG 420
Qy      156  ValMetCysVal1leuLeuTrpAla1leuSerLeuLeuAsnSer11eLeuGluTrpMetPhe 175
Db      421  GTGCTGTGTCTCTGCTCTGCGGCGCTCTCTCTGCTGTATGATGATGATGATGATGATG 480
Qy      176  CysAspPheLeuPheSerGlyAlaAspSerVal1TrpCysGluThrSerAspPhe11eThr 195
Db      481  TGTGACTTCTGTTAGTGTGCTGATTTAGTTGAGTGAAGCTCAGATTTCATCCCA 540
Qy      196  11eAla1TrpLeuVal1PheLeuCysVal1leuCysGlySerSerLeuVal1leuVal 215
Db      541  GTGCTGTGCTGATTTTATATGATGATCTCTGATTTTCCAGCCCTGCTGCTGCTGCTC 600
Qy      216  Arg11eUeuCysGlySerArgLysMetProLeuThrArgLeuTyVal1Thr1leUeu 235
Db      601  AGATCTCTGTGTGGATCCCGAAGATGCCGCTGACCAAGGCTGATGATGATGATGATGATG 660
Qy      236  ThrVal1euVal1Phe1euLeuCysGlyLeuProPheGly11egIntPala1euPheSer 255
Db      661  ACAGTGTGCTCTCTCTCTCTGCGGCTCTCTCTCGGCATTTCCGAGGCGCTTAATTAC 720
Qy      256  Arg11eHis1euAspTrpLysVal1euPheCysHisVal1His1euValSer11ePheLeu 275
Db      721  AGAGTGCACCTGATTTTGAAGTCTTAATGTATGATATGATGATGATGATGATGATGATG 780
Qy      276  Ser1Ala1euAsnSerSer1AlaAsnPro11e1eTyPhePheVal1GlySerPheArg1n 295
Db      781  TCTCTCTTAACAGTATGTCACCAACCCCATCATTTACTTCTTCTGAGGCTCTTTAGGCGAG 840
Qy      296  ArgGlnAsnArgGlnAsnLeuLysLeuVal1euGlnArgAla1euGlnAspThrProGlu 315
Db      841  CGTCAAAATAGCGAAGAACTGAAAGCTGTTCTCCAAAGGCGCTGCGAGCAAGCTGAG 900
Qy      316  ValAspGlnGlyGlyTyTrpLeuProGlnGluThrLeuGlnLeuSerGlySerArgLeu 335
Db      901  GTGGATTAAGGTGAAGGCGAGCTTCTCGAGAAAGCCTGAGAGCTGTGCGAAGGAGATTG 960

```

RESULT 7

```

US-09-254-227A-2
; Sequence 2, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: rat
US-09-254-227A-2

```

Alignment Scores:

Pred. No.:	1.99e-75	Length:	1011
Score:	834.50	Matches:	174
Percent Similarity:	68.69%	Conservative:	52
Best Local Similarity:	52.89%	Mismatches:	92
Query Match:	47.33%	Indels:	11
DB:	4	Gaps:	6

US-09-867-570-2 (1-337) x US-09-254-227A-2 (1-1011)

```

Qy 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAengly 32
Db 34 TTTGGACATGATGATCCACCATCTCATCCCTCAGACAGAACTCTCAACACGATATAA 93
Qy 33 ArgGluGluThrProCysThrYsglnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 94 ACTGGTCATCCAGTTGC---AGGCCAATCTCACCCCTGCTCTGCTGCCCATCATC 150
Qy 53 SerLeuValAlaLeuThrGlyValMetAlaValLeuThrLeuGlyCysArgMetArg 72
Db 151 ACCGCTTGATGATGACAGAAACACCATGTGACTGTGCTTGGATTCGCGATGCGC 210
Qy 73 ArgAsnAlaValSerIleThrLeuAsnLeuValAlaAspPheLeuPheLeuSer 92
Db 211 AGMAAGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 270
Qy 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIle-----ArgHisPro 108
Db 271 TGCCATTTATTTATGATCTCTGATGCGGATCATGATGATGATGATGATGATGATG 330
Qy 109 IleSerIys---IleLeuSerProValMetThrPheProTyPheIleGlyLeuSerMet 127
Db 331 TTAAGCAAGAAATCTTAGGCAATGATGATGATGATGATGATGATGATGATGATG 390
Qy 128 LeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleThrIleCys 147
Db 391 CTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 450
Qy 148 ArgArgProArgTyIleuSerSerValMetCysValLeuLeuThrAlaLeuSerLeu 167
Db 451 CACCCCAAGAAATCTTAGGCAATGATGATGATGATGATGATGATGATGATGATG 510
Qy 168 ArgSerIleLeuGluThrMetPheCysAspPheLeuPheSerGlyAlaAspSerVal 187
Db 511 ATGGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
Qy 188 CysGluThrSerAspPheIleThrIleAlaThrLeuValPheLeuCysValValLeu 207
Db 571 ---AAAATGTTGATTTATTTGATGATGATGATGATGATGATGATGATGATG 627
Qy 208 GlySerSerLeuValLeuValAlaGlyIleCysGlySerArgIleMetProLeuThr 227
Db 628 GGGTCAGTCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 687
Qy 228 ArgLeuTyValThrIleLeuLeuThrValLeuValPheLeuLeuGlyLeuProPhe 247
Db 688 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 747
Qy 248 GlyIleGlnThrAlaLeu-----PheSerArgIleHisLeuAspThrIleVal 265
Db 748 GGGCTTACTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 801
Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAspPro 285
Db 802 TGTCACTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATG 861
Qy 286 IleTyPhePheValGlySerPheArgGlnArgGlnAsnArgIleAsnLeuVal 305
Db 862 ATTACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyGlyThrLeuProGln 325

```

DB 922 CTTAAAGGCTCTGAGAGACTCTGAGAGATGATATATACAGACGACCATGTTGAG 981

Qy 326 GluThrLeuGluLeuSerGlySerArg 334

DB 982 AAACCCAGATGATCTCGAAGAGAGA 1008

RESULT 8

US-09-495-050A-164

/ Sequence 164, Application US/09495050A

/ Patent No. 6492505

/ GENERAL INFORMATION:

/ APPLICANT: Roopa, Reddy

/ APPLICANT: Guegler, Karl, J.

/ APPLICANT: Au-Young, Janice

/ TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P

/ FILE REFERENCE: PA-0013 US

/ CURRENT APPLICATION NUMBER: US/09/495,050A

/ PRIOR FILING DATE: 2000-01-31

/ PRIOR APPLICATION NUMBER: 60/118,318

/ PRIOR FILING DATE: February 1, 1999

/ NUMBER OF SEQ ID NOS: 305

/ SOFTWARE: PERL Program

/ SEQ ID NO 164

/ LENGTH: 291

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: misc. feature

/ OTHER INFORMATION: Incyte ID No. 6492505 1909132CT1

US-09-495-050A-164

Alignment Scores:

Pred. No.:	3.47e-45	Length:	291
Score:	529.00 <td>Matches:</td> <td>97 </td>	Matches:	97
Percent Similarity:	100.00% <td>Conservative:</td> <td>0 </td>	Conservative:	0
Best Local Similarity:	100.00% <td>Mismatches:</td> <td>0 </td>	Mismatches:	0
Query Match:	30.01% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	4	Gaps:	0

US-09-867-570-2 (1-337) x US-09-495-050A-164 (1-291)

```

Qy 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyPheIleGlyLeu 125
Db 1 CGGCATCCCATCTCAAAATCTCACTGATGATGATGATGATGATGATGATGATG 60
Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleThr 145
Db 61 AGCATGCTAGGCGCATGACGACGAGCGCTGCTGCTCATCTGCGCCCATCTGTG 120
Qy 146 HisCysArgArgProArgTyIleuSerSerValMetCysValLeuLeuThrAlaLeu 165
Db 121 CACTGCGCGCGCCGCAAGATCTGATGATGATGATGATGATGATGATGATGATG 180
Qy 166 LeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPheSerGlyAlaAsp 185
Db 181 CGCTGCGGAGATCTCTGAGATGATGATGATGATGATGATGATGATGATGATG 240
Qy 186 ValThrCysGluThrSerAspPheIleThrIleAlaThrLeuValPheLeu 202
Db 241 GTTGTGTGAAACGTCAGATTTCAATTCATGCGGTGCTGCTGCTGCTGCTGCT 291

```

RESULT 9

US-09-016-434-330

/ Sequence 330, Application US/09016434

/ Patent No. 6500938

/ GENERAL INFORMATION:

/ APPLICANT: Janice Au-Young

/ APPLICANT: Jeffrey J. Selhamer

/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

/ TITLE OF INVENTION: PATHWAY GENE EXPRESSION

/ NUMBER OF SEQUENCES: 1490

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

RESULT10
US-09-016-434-1264
; Sequence 1264, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhammer

QY 51 lvalaSerleuValalaleThrcylYAsnalavalValleuTrpIleuLeuGlyCYsArg 70
Db 1246 CTGGCTTCCTCTGTGGGGCTTTATTTGAATGGCACTGCTTCTGCGCTTGTCTGGG 1305
QY 71 MetArgArgAnaIaValSerIleTyrIleuAsnIleuValaIaAlaAspPheIeuPhe 90
Db 1306 ---GCCAGCAATCCCTCATGCTATACATCTCCACGCGGTGGTGGACGTATCAT 1366
QY 91 LeuSerGlyHisIleIleCYSerProIeuArgLeuIleAsnIleArg----- 106
Db 1363 CTT-----TGCTGCTCGGACGATGGGGTTCTTACAGCGACCTTGCTAACTAT 1410
QY 107 HisProIle-----SerValIleuSerProValMetThrPhe 119
Db 1411 CATGAGACGCGTTTTTATTCCTGATTTCTCGGCCCATATGTCTCC-----TTC 1461
QY 120 ProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrgArgCYsLeuSerIle 139
Db 1462 TCCTTTGAAGGTGCTCTGTCTGCTGTCGTCATGACACAGACGGGTGTGTGATC 1521
QY 140 LeuTrpProIleTrpTyrHisCYsArgArgProArgTyrLeuSerSerValMetCYsVal 159

```

Db 1522 CTCCTCCCATCTGGATGACAGATCCACCGCCCAAAATACATCTAATGTGTGTGACCC 1581
Qy 160 LeuLeuTriAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeu 179
Db 1582 CTCATCTGGGGCTTCCTTTTTCATCAACATAGTAATAATCACTTTCTCACTTAC 1638
Qy 180 PheSerGlyAlaAspSerValTrpCysGluThrSerAsp-----PheIleThrIle 196
Db 1639 -----TGGAAACATGTAAGGACATGTGTCTATTTCTTAAGCTT 1677
Qy 197 AlaTrpLeu-----ValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeu 214
Db 1678 TCTGGGCTCTTCATGATCTTCTTCACTTGATGTGTGTGTGAGTCTGACTTAC 1737
Qy 215 ValAlaGlyLeuCysGlySerArgGlySerMetProLeuThrArgLeuTrpValThrIleLeu 234
Db 1738 ATTAAATTCCTGTGTGTCTCCACGACGAAAGGCCACCAAGGCTTATGGGTGTGACG 1797
Qy 235 LeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPhe 254
Db 1798 ATCTGGCCCCCATGTCTCTACTCTTGAGCCCTTACCCCTGAGCGTG----- 1842
Qy 255 SerArgIleHisLeuAspTrpIleValLeuPheCysHisValHisLeuValSerIlePhe 274
Db 1843 GCACCCCTCATTAACAGATTTCAAAATGTGTGTGCACACCTCTATTTAATTTCTTGTTC 1902
Qy 275 LeuSerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGlySerPheArg 294
Db 1903 CTC--ATTATTAACAGACGACGCCCAACCTTATCATTTATTTCTTTTGGGACCTCAAG 1959
Qy 295 GluArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsnPro 314
Db 1960 AAGAAAGGCTGAAGGAATCTCTCAGAGTGATTTCTCAACGGCGTTAGCATTAAGCCA 2019
Qy 315 GluVal 316
Db 2020 GAGGTG 2025

RESULT 11
US-09-016-434-1225
; Sequence 1225, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1388 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9187388
; US-09-016-434-1225

Alignment Scores:
Pred. No.: 3,35e-34 Length: 1388
Score: 430.50 Matches: 103
Percent Similarity: 58.46% Conservative: 56
Best Local Similarity: 37.87% Mismatches: 94
Query Match: 24,42% Indels: 19
DB: Gaps: 8

US-09-867-570-2 (1-337) x US-09-016-434-1225 (1-1388)
Qy 52 ValSerLeuValAlaLeuThrGlyAsnAlaValAlaLeuTrpLeuGlyCysArgMet 71
Db 391 ATCTCCCAAGTGGGGGTTGTGAGATGGGATTTCTCTGTGCTTCTGCTTCCGATG 450
Qy 72 ArgArgAsnAlaValSerIleTrpIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 451 AGAAGAAATCCCTTCACTGTCTTACACCCACTGTCTTATGCGACACTCTCACTGCTC 510
Qy 92 SerGlyHisIleIleCysSer-----ProLeuArgLeuIleAsnIleArg 106
Db 511 TTCTGATTTTCACTTGTGTCTTATCGACTGTGCTTAGATTAAGACTTTCTTCCGCA 570
Qy 107 HisProIleSerIleLeuSerProValMetThrPheProIleGlyLeuSer 126
Db 571 TACTACACATTTGTCACTTATCACTGATCACTTCTTGTGTGCTACACAGGCTCTAT 630
Qy 127 MetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpHis 146
Db 631 CTGCTGACGCGCATTAAGTGTGAGAGGTGCTGTGACGCTTATACCCATCTGTACCGA 690
Qy 147 CysArgArgProArgTrpLeuSerSerValMetCysValLeuLeuTriAlaLeuSerLeu 166
Db 691 TGCATCGCCCAAGTACAGATGCGGATGTGTGTGCTTGTGGCTTTCTTCTTGC 750
Qy 167 LeuArgSerIleLeuGluTrpMetPheCys-----AspPheLeuAsnSerGly 182
Db 751 TTGGTGACCAACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
Qy 183 AlaAspSerValTrpCysGlu---ThrSerAspPheIleThrIle---AlaTrpLeuVal 200
Db 811 AATGAC-----TGCAGACAGCATCATCTTTATAGCCATCTGAGCTTGTGTC 861
Qy 201 PheLeuCysValValLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
Db 862 TTACGCCCTTCACTGCTGTG---TTCACACACATTTGTGTGTGTGTGTGTGTGTGTGT 918
Qy 221 SerArgIleMetProLeuThrArgLeuTrpValThrIleLeuLeuThrValLeuValPhe 240
Db 919 AGCTGGGCTTCCATCTTCCTCAAGCTTTCATAGTATGATGATGATGATGATGATGATG 978
Qy 241 LeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 979 CTCATCTTGCATAGCCCATGAGACTCTTATCTGCTGTATGAGAT----- 1029
Qy 261 TrpIleValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 1030 TGGTCAGC---TTGGGAACCTTACACACATTTCTGCTTCTTCTTCCATTAACACT 1086
Qy 281 SerAlaAsnProIleIleTrpPhePheValGlySerPheArgGlnArgGlnAsnArgGln 300

```

Db 1087 AGGCGCAACCTTCTTCTTCTTGTGGAGACGATAGAGAGAGATTCAGAG 1146
QY 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1147 TCCTTAAAGTTGTTGACACGAGGCTTTCAAAGAT 1182

RESULT 12
PCT-US93-06251-26
Sequence 26, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1388 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-26

Alignment Scores:
Pred. No.: 3,35e-34 Length: 1388
Score: 430.50 Matches: 103
Percent Similarity: 58.46% Conservative: 56
Best Local Similarity: 37.87% Mismatches: 94
Query Match: 24.42% Indels: 19
DB: Gaps: 8

US-09-867-570-2 (1-337) x PCT-US93-06251-26 (1-1388)

QY 52 ValSerLeuValAlaLeuThrGlyAsnAlaValLeuTrpLeuLeuGlyCysArgMet 71
Db 391 ATCTCCCGAGTGGGGTTGTGGAATGGGATTCCTCTGGTTCCTGCTCCGATG 450
QY 72 ArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 451 AGAAGAAATCCCTTCACTGCTACATCACCACCTGTCTATCGCAGACATCTCACTCTC 510
QY 92 SerGlyHisIleLeuCysSer-----ProLeuArgLeuIleAsnIleArg 106
Db 511 TTCTGATTTTCACTTGTCTATCGATAGATGCTTAATATAGCTTCTTCCGCGAT 570
QY 107 HisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeuSer 126
Db 571 TACTACACATTTGTCATCATATATGAGTCTTTCTGTTGGCTTACACAGCGGCTCTAT 630

QY 127 MetLeuSerAlaIleSerThrGlnArgCysLeuSerIleLeuTrpProIleTyrHis 146
Db 631 CTGCTGACGGCCATTAAGTGGAGAGGCTGTCAGTCTTATCCCATCTGTATCCGA 690
QY 147 CysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeu 166
Db 691 TGCATGCGCCCAAGTACAGTCCAGTGGCATTTGGTCTGTCCCTTCTGTGGCTTTCTTTC 750
QY 167 LeuArgSerIleLeuGlnTrpMetPheCys-----AspPheLeuPheSerGly 182
Db 751 TTGTGTACACACATGAGTATGTCATGTCATGCAGACAGAGAGAGAGATCACTTCGG 810
QY 183 AlaAspSerValTrpCysGlu---ThSerAspPheIleThrIle---AlaTrpLeuVal 200
Db 811 AATGAC-----TCCAGACAGTCATCATCTTATACCATCTGAGGCTTCCTGGTC 861
QY 201 PheLeuCysValValLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
Db 862 TTCACGCCCCCTCATGCTGTG---TCCAGACCATCTTGTGCTGGAAGATCCGAGAAC 918
QY 221 SerArgLysMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 240
Db 919 ACGTGGGCTTCCCATCTCTCCAGCTTATAGTCATGTCATGTCACCATCATATATTC 978
QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 979 CTATCTTCGCTAGCCCATGAGACTCTTACCTGCTGTACATATGAGTAT----- 1029
QY 261 TrpLysValLeuPheCysHisValIleLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 1030 TGTGTGAC---TTTGGAACTTACACCATTTCCCTGCTTCTTCCACATATACAGT 1086
QY 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGln 300
Db 1087 AGGCGCAACCTTCACTTACTTCTTGTGGAGAGAGTAAAGAGAGATTCAGAG 1146
QY 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1147 TCCTTAAAGTTGTTGACACGAGGCTTTCAAAGAT 1182

RESULT 13
5320941-1
Patent No. 5320941
Applicant: Young, Dallan, Wigler, Michael H., Passano
Otlavio
TITLE OF INVENTION: DNA SEQUENCES ENCODING MMS ONHCOGENE,
POLYPEPTIDES ENCODED THEREFROM AND DIAGNOSTIC AND OTHER METHODS
BASED THEREFROM
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/872,087
FILING DATE: 06-JUN-1986
SEQ ID NO: 1:
LENGTH: 1327
5320941-1

Alignment Scores:
Pred. No.: 2.27e-30 Length: 1327
Score: 392.50 Matches: 97
Percent Similarity: 55.20% Conservative: 52
Best Local Similarity: 34.77% Mismatches: 72
Query Match: 22.26% Indels: 53
DB: Gaps: 9

US-09-867-570-2 (1-337) x 5320941-1 (1-1327)

QY 52 ValSerLeuValAlaLeuThrGlyAsnAlaValLeuTrpLeuLeuGlyCysArgMet 71
Db 390 ATCTCCCGAGTGGGGTTGTGGAATGGGATTCCTCTGGTTCCTGCTCCGATG 449
QY 72 ArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 450 AGAAGAAATCCCTTCACTGCTACATCACCACCTGTCTATCGCAGACATCTCACTCTC 509

```

QY 92 SerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIys 111
Db 510 -----TTCGTG-----ATTTCACTGCTGTATGACATATGCTTAAATAT 551
QY 112 IleLeuSer-----ProValMetThr-----PhePro 120
Db 552 GAGCTTTTCTTGCGCATCTACACAAATGTCACACTTATCACTAGAGTTTCTGTTGGC 611
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
Db 612 TACAACACGGGCTCTATCTGCTGACGCGCATTAAGTGAGAGGCTGCTGACAGCTT 671
QY 141 TrpProIleTrpTyrHisArgArgProArgTyrLeuSerSerValMetCysValLeu 160
Db 672 TACCCCATCTGTATCCCATGCGCCATGCGCCCAAGTACAGTACAGTGGTGTGGCTT 731
QY 161 LeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeu 180
Db 732 CTGTGGGCTCTTTCTTGCTGTGACACCATGAGATATGTCATGTGCAACACAGAA 791
QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAsp-----PheIle 194
Db 792 GAAGAGAGTGAAGTCT-----CCGAATGACTGCCACAGCTCATCATCTTATA 839
QY 195 ThrIle---AlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeu 213
Db 840 GCCATCTGAGGCTCTGCTGTCTTCAAGGCTCATCTGCTGCTG---TCCAGCACCATCTG 896
QY 214 LeuValArgIleLeuCysGlySerArgIysMetProLeuThrArgLeuTyrValIle 233
Db 897 GTCCGGAAGATCCCAAGAAACAGTGGGCTTCCCATCTCCCAAGTTTACATAGCATC 956
QY 234 LeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeu 253
Db 957 ATGGTCACCATCATATATCTCATCTTCGCTATGCGCATGAGACTGCTTTCCTGCTG 1016
QY 254 PheSerArgIleHisLeuAspTrpTyrValLeuPheCysHisValHisLeuValSerIle 273
Db 1017 TAC----- 1019
QY 274 PheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPhe 293
Db 1020 -----TACAGTAGCGCCCAACCTTCTTACTTCTTGTGGAGAGAGAT 1064
QY 294 ArgGlnArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1065 AAGAAGAGAGATTCAGACAGTGTAAAGTTGCTGACACAGGCTTTCAAGAT 1121

```

RESULT 14

```

US-09-484-970B-134
/ Sequence 134, Application US/09484970B
/ Patent No. 6426186
/ GENERAL INFORMATION:
/ APPLICANT: Jones, Karen A.
/ APPLICANT: Volkmuth, Wayne
/ APPLICANT: Walker, Michael G.
/ TITLE OF INVENTION: BONE REMODELING GENES
/ FILE REFERENCE: PB-0014 US
/ CURRENT APPLICATION NUMBER: US/09/484, 970B
/ CURRENT FILING DATE: 2000-01-18
/ NUMBER OF SEQ ID NOS: 172
/ SOFTWARE: PERL Program
/ SEQ ID NO 134
/ LENGTH: 2435
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6426186 247789.2CBI
/ LOCATION: 93, 128, 132, 143-144, 2419, 2427, 2429
/ OTHER INFORMATION: a, t, c, g, or other

```

US-09-484-970B-134

```

Alignment Scores:
Pred. No.: 7,15e-29 Length: 2435
Score: 381.50 Matches: 100
Percent Similarity: 52.26% Conservative: 50
Best Local Similarity: 34.84% Mismatches: 112
Query Match: 21.64% Indels: 25
DB: 4 Gaps: 7

```

US-09-867-570-2 (1-337) x US-09-484-970B-134 (1-2435)

```

QY 48 LeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValLeuTrpLeuLeu 67
Db 642 CTCCTCTGCTGTGGCTGTG-----GCCAAGGGCTGTGCTCTGCTGTTTTC 692
QY 68 GlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAsp 87
Db 693 GGCCTTCCATCAAGAGAACCCCTTCTCCATCTTCTTCGACCTGGCCAGCGCAT 752
QY 88 PheLeuPheLeuSerGlyHisIleIleCysSerProLeuArg----- 101
Db 753 GTGGGCTACTCTTTCAGCAAGGCGGTCTTCATCTGAAACAGGGGGCTTCTGGGC 812
QY 102 ---LeuIleAsnIleArgHisProIleSerIysIleLeuSerProValMetThrPhePro 120
Db 813 AGCTTGGCCGATCAATCCGAGCGGTGCGCGGCTGTGGGCTGCGATG----- 863
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeu-SerIleLe 140
Db 864 TTCCTTACCGCGGTGAGCTCTGCGCGCGGCTCGACCTCGAGCGCTGCGCTTGGTAT 923
QY 140 uTrpProIleTrpTyrHisCys-ArgArgProArgTyrLeuSerSerValMetCysVal 160
Db 924 CTTCCTCCGCGGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCTGCGCGCC 983
QY 160 euLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeu 180
Db 984 TGCTGTGGGTCTGTGCTCTCTGTCGTGATCCTGCTGCAACATCTTGGCGTCTGCG 1043
QY 180 heSerGlyAlaAspSerValTrpCysGluThrSerAsp---PheIleThrIleAlaTrp 199
Db 1044 GCCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 1103
QY 199 euValPheLeuCysValValLeuCysGlySerSerValLeuLeuValArgIleLeu 219
Db 1104 TCTGCTCTGCTGCGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
QY 219 yseGlySerArgIysMetProLeuThr---ArgLeuTyrValThrIleLeuLeuVal 238
Db 1164 GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1223
QY 238 euValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIle 258
Db 1224 CCGCTCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
QY 258 IseLeuAspTrp-----LysValLeuPheCysHisValHisLeuValSerIlePheLeu 276
Db 1274 -----TGGGCTTCCAGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCA 1325
QY 276 exAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 296
Db 1326 TCGCATCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1385
QY 296 rseGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsnTrpProGlu 316
Db 1386 AGCGGCTGTGGAGCGCGCTCAGGCTGCTTCCAGCGGCGCGCTGCGGAGCGGCTGAGC 1445
QY 316 AlaAspGluGlyGly 321
Db 1446 TGGGGAGGCGCGGGGCG 1462

```

RESULT 15

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: December 1, 2004, 22:50:53 ; Search time 553 Seconds

(without alignments)
3348.557 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
1 MESKSSWYIRLGLFLSHDSTI.....EGGGWLPQETLELSGRLEQ 337

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Xgapext 0.5	
Egapop 6.0 , Egapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3694831 seqs, 274740616 residues

Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xip
-Q/cgnt2 1/USPTO.spool.p/US09867570/runat.01122004.161124.6393/app.query.fasta_1.519
-DB=Published Applications NA -QFMT=fasta -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITEXT=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=JOCL -OUTFMT=ptc -NORM=ext -HRAPIZ=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MMAP -LARGESUBRY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YCAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Published Applications NA:
1: /cgnt2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgnt2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgnt2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgnt2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgnt2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgnt2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgnt2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgnt2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgnt2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgnt2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgnt2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgnt2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgnt2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgnt2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgnt2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgnt2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	2618	11 US-09-867-570-1	Sequence 1, Appl1
2	1727.5	98.0	8622	11 US-09-867-570-3	Sequence 3, Appl1
3	1722.5	97.7	1369	15 US-10-292-798-1273	Sequence 1273, Ap
4	1721	97.6	1400	14 US-10-183-116-30	Sequence 30, Appl
5	1721	97.6	1400	15 US-10-225-567A-673	Sequence 673, App
6	1688	95.7	969	9 US-09-995-225-19	Sequence 19, Appl
7	1688	95.7	969	10 US-09-995-225-19	Sequence 19, Appl
8	1683	95.5	969	15 US-10-401-397A-1	Sequence 1, Appl1
9	1663	94.3	909	15 US-10-337-467-9	Sequence 9, Appl1
10	1557	88.3	909	15 US-10-337-467-9	Sequence 9, Appl1
11	1416	80.3	1369	15 US-10-017-161-1055	Sequence 1055, Ap
12	1416	80.3	1369	15 US-10-292-798-897	Sequence 897, App
13	1416	80.3	1370	14 US-10-183-116-15	Sequence 15, Appl
14	1402.5	79.6	1370	15 US-10-017-161-1599	Sequence 1599, Ap
15	1397	79.2	997	16 US-10-072-012-171	Sequence 171, App
16	1390	78.8	1369	15 US-10-292-798-1041	Sequence 1041, Ap
17	1388	78.7	1604	14 US-10-183-116-32	Sequence 32, Appl
18	1388	78.7	1604	15 US-10-225-567A-668	Sequence 668, App
19	1383	78.4	969	15 US-10-079-384-3	Sequence 3, Appl1
20	1383	78.4	969	15 US-10-240-998-3	Sequence 3, Appl1
21	1383	78.4	969	15 US-10-321-807-19	Sequence 19, Appl
22	1383	78.4	969	15 US-10-237-467-11	Sequence 11, Appl
23	1383	78.4	969	16 US-10-343-650A-43	Sequence 43, Appl
24	1383	78.4	969	17 US-10-321-807-19	Sequence 19, Appl
25	1383	78.4	969	17 US-10-314-048A-19	Sequence 19, Appl
26	1383	78.4	969	18 US-10-481-161-1	Sequence 1, Appl1
27	1367	77.5	969	9 US-09-995-225-17	Sequence 17, Appl
28	1367	77.5	969	10 US-09-995-225-17	Sequence 17, Appl
29	1367	77.5	969	15 US-10-237-467-3	Sequence 3, Appl1
30	1365	77.4	1005	16 US-10-072-012-173	Sequence 173, App
31	1339.5	75.4	1005	16 US-10-072-012-177	Sequence 177, App
32	1293	73.3	994	16 US-10-072-012-169	Sequence 169, App
33	1279	72.5	1997	14 US-10-219-834-7	Sequence 7, Appl1
34	1270	72.0	966	15 US-10-240-998-6	Sequence 6, Appl1
35	1265	71.8	966	15 US-10-240-998-9	Sequence 9, Appl1
36	1227	69.6	793	15 US-10-101-510-239	Sequence 239, App
37	1031.5	58.5	993	17 US-10-467-616-1	Sequence 1, Appl1
38	1009.5	57.3	1018	16 US-10-311-671-26	Sequence 26, Appl
39	1009.5	57.3	1163	9 US-09-750-373-11	Sequence 11, Appl
40	1009.5	57.3	1176	18 US-10-481-161-11	Sequence 11, Appl
41	1009.5	57.3	1300	14 US-10-183-116-17	Sequence 17, Appl
42	1009.5	57.3	1300	15 US-10-225-567A-648	Sequence 648, App
43	1009.5	57.3	1393	15 US-10-017-161-1053	Sequence 1053, App
44	1009.5	57.3	1393	15 US-10-292-798-895	Sequence 895, App
45	1009.5	57.3	1770	9 US-09-920-068A-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-867-570-1
Sequence 1, Application US/09867570
Publication No. US20040076951A1
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: C1000900-CIP
CURRENT APPLICATION NUMBER: US/09/867,570
PRIOR APPLICATION NUMBER: 2001-05-31
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2618
TYPE: DNA
ORGANISM: Human
US-09-867-570-1

Alignment Scores:
 Pred. No.: 1.59e-175 Length: 2618
 Score: 1763.00 Matches: 337
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 11 Gaps: 0

US-09-867-570-2 (1-337) x US-09-867-570-1 (1-2618)

```

QY      1 MetGluSerLysSerSerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIle 20
DB      447 ATGGAAGCAAAATCTCTCATGGCTCATCAACCTGGGTTTCTAGCATGATTCACCAATC 506
QY      21 ProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGluThrProCysTyrLys 40
DB      507 CCAGCTCTGGGTACAGAACTGACCAATCAACGAGGAGGAGACTCCTTGCTCAAG 566
QY      41 GlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsn 60
DB      567 CAGACCTGAGCTTACAGGGGCTGAGGTGATCGTTCCCTGTGCGCTGACAGAAAC 626
QY      61 AlaValIleLeuTrpLeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIle 80
DB      627 GCGGTGTGCTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
QY      81 LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeu 100
DB      687 CTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
QY      101 ArgLeuIleAsnIleArgHisProIleSerLysIleLeuSerProValMetThrPhePro 120
DB      747 GCGCTCATCAATATCCGATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 806
QY      121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
DB      807 TACTTTATAGGCTTAAGCATGCTGAGCGCATGACAGACCGAGGCTGCTGCTGCTGCTGCTGCT 866
QY      141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
DB      867 TGGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
QY      161 LeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPhe 180
DB      927 CTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
QY      181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
DB      987 AGTGTGCTGATTTCTGTTGGTGTGAAGCTCAGATTTCATTACATTCGGTGGCTGCTGCTGCT 1046
QY      201 PheLeuCysValValLeuCysGlySerSerLeuValLeuValAlaArgIleLeuCysGly 220
DB      1047 TTTTATATGTTGTTCTCTGTGGGTGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
QY      221 SerArgLysMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 240
DB      1107 TCCCGAAGATGCCGCTGACAGGCTGACGATGACCAATCTCTCAACAGCTGCTGCTGCTGCT 1166
QY      241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
DB      1167 CTCCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
QY      261 TrpLysValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
DB      1227 TGGAAAGCTTATTTTGTGATGTCATGATGTTCCATTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1286
QY      281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGlnArgGln 300
DB      1287 AGTGCACACCCATCATCTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346
QY      301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320
  
```

```

DB      1347 AACCTGAAGCTGTTCTCCAGAGGGCTCTGACAGACAGCCCTGAGTGGATGAAGTGA 1406
QY      321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
DB      1407 GGGTGGCTTCTCCAGGAAACCTGAGGCTGTGGAGAGCAATGAGTGAAGCAG 1457
  
```

RESULT 2
 US-09-867-570-3
 ; Sequence 3, Application US/09867570
 ; Publication No. US20040076951A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: CL000900-CIP
 ; CURRENT APPLICATION NUMBER: US/09/867,570
 ; PRIOR FILING DATE: 2001-05-31
 ; PRIOR APPLICATION NUMBER: 09/695,045
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 8622
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-867-570-3

Alignment Scores:
 Pred. No.: 5.4e-171 Length: 8622
 Score: 1727.50 Matches: 332
 Percent Similarity: 96.81% Conservative: 2
 Best Local Similarity: 96.23% Mismatches: 9
 Query Match: 97.99% Indels: 1
 DB: 11 Gaps: 1

US-09-867-570-2 (1-337) x US-09-867-570-3 (1-8622)

```

QY      2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12
DB      7459 CAGAGATCAAAACAGCTGGATGATCATCTGTTCTGTTTCCAGGGGTATCAGACTGGAG 7518
QY      13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
DB      7519 TTTCTAGCATGATTCATCAACATCCAGCTTGGGTACGAACTGACACCAATCAACGGA 7578
QY      33 ArgGluGluThrProCysTyrLysGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
DB      7579 CGTGAAGAGATCTCTGCTTACAAAGCAGACCTGAGCTTCAAGGGGCTGACGTGCTGCTGCT 7638
QY      53 SerLeuValAlaLeuThrGlyAsnAlaValIleLeuTrpLeuGlyCysArgMetArg 72
DB      7639 TCCCTTGTGCGCTGACAGAAACGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7698
QY      73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSer 92
DB      7699 AGGAAGCTGTCTTCATCTCAATCTCAACCTGGTGGGCGGAGCTTCTTCTTCTTAC 7758
QY      93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerLysIle 112
DB      7759 GGGCAATTATATGTTGCGCGTTAGCGCTCATCAATATCCGCAATCCCATCCCAAAATC 7818
QY      113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
DB      7819 CTCAGTCTGTGATGATGCTTCCCTTCACTTATAGGCTTAAGATCTGATGAGCCCATCAGC 7878
QY      133 ThrGluArgCysLeuSerIleLeuTrpProIleTyrHisCysArgArgProArgTyr 152
DB      7879 ACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7938
QY      153 LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGlu 172
DB      7939 CTGTCACTGGTATGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7998
  
```

Qy 173 TrpMetPheCysAbpPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
 Db 7999 TGGATGTTCTGTGACCTTCGTTTAGTGCGTCGATTCGTTGGTGTGAACCGTCAGAT 8058
 Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuVal 212
 Db 8059 TTCATTACAAACGGGCGCTGGGTTTATTATGTTGTTCTCTGTGGGTCACAGCTGGTC 8118
 Qy 213 LeuLeuValAlaArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThr 232
 Db 8119 CTGCTGTCAGGANTCTCTGTGATCCCGAAGATGCGCGACCAAGCTGTAGTACC 8178
 Qy 233 IleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAla 252
 Db 8179 ATCCTCTCACAGCTGCTGCTCTCTCTGTCGCTGCGCTTGGCATTCAGTGGGCC 8238
 Qy 253 LeuPheSerArgIleHisLeuAspTrpValLeuPheCysHisValHisLeuValSer 272
 Db 8239 CTGTTTCCAGATCCACCTGGATTTGAAGTCTTATTGTTCATGTGCATCTAGTTTCC 8298
 Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValAlGlySer 292
 Db 8299 ATTTCCTGTCCTCTCTTAACAGACAGTCCCAACCCCATTTACTTCTGTCGGCTCC 8358
 Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuValLeuGlnArgAlaLeuGlnAsp 312
 Db 8359 TTTAGGCAAGCGTCAAAATAGGCAAGACCTGAAGCTGTTCTCCAGAGGGCTTCAGAGAC 8418
 Qy 313 ThrProGluValAlaArgGluGlyGlyIleTyrLeuProGlnGluThrLeuGluLeuSerGly 332
 Db 8419 ACGCCTAGAGTGAAGTGAAGTGAGGGTGGCTTCTCAGAGAAACCTGAGACTGTCCGGA 8478
 Qy 333 SerArgLeuGluGln 337
 Db 8479 AGCAGATTGGAGCAG 8493

RESULT 3

US-10-292-798-1273
 ; Sequence 1273, Application US/10292798
 ; Publication No. US2003023583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUMA, MAKIKO
 ; APPLICANT: ASAI, KIYOSHI
 ; APPLICANT: AKIYAMA, YUTAKA
 ; APPLICANT: AUBURATANI, HIROYUKI
 ; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
 ; FILE REFERENCE: 084335/166
 ; CURRENT APPLICATION NUMBER: US/10/292, 798
 ; PRIOR FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: 10/017,161
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: JP 2001-246789
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 2070
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1273
 ; LENGTH: 1369
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; LOCATION: source
 ; FEATURE: (1)..(1369)
 ; NAME/KEY: CDS
 ; LOCATION: (201)..(1169)
 ; US-10-292-798-1273

Alignment Scores: 1,14e-171 Length: 1369
 Score: 1722.50 Matches: 331
 Percent Similarity: 96.81% Conservative: 3

Best Local Similarity: 95.94% Mismatches: 2
 Query Match: 97.70% Indels: 9
 DB: 15 Gaps: 1
 US-09-867-570-2 (1-337) x US-10-292-798-1273 (1-1369)
 Qy 2 GluSerIysSerSerTrp-----ValIleArgLeuGly 12
 Db 132 CAGAGATCAAAACAGCTGGTATCATCTGCTTTCAGAGTCATCAGCTGAG 191
 Qy 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
 Db 192 TTTCTGAGCATGATTTCAACATCCAGCTTGGGTACAACTGACCAATCAAGGA 251
 Qy 33 ArgGluGluThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
 Db 252 CGTGAGAGACTCCTTCTCAAGCAGACCTGAGCTTCAAGGGGCTGAGCGTCATCGTT 311
 Qy 53 SerLeuValAlaLeuThrGlyAsnAlaValValLeuThrLeuLeuGlyCysArgMetArg 72
 Db 312 TCCCTTGTCCGCTGACAGAGAAACGCGTGTGTCTGTGGCTCTCCGAGCTCCGATTCGC 371
 Qy 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSer 92
 Db 372 AGAAGCCTGTCTTCATCTACATCTCAACCTGGTCGCGGCCACCTTCTTCTTACG 431
 Qy 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
 Db 432 GGCCACTATATAGTTGGCCGTTACGCTCATCAATCCGCCATCTCCAAATC 491
 Qy 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
 Db 492 CTCAGTCTGTGATGACCTTTCCTTATATAGCCTTAAGCATGATGAGGCGCATCAGC 551
 Qy 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyr 152
 Db 552 ACCAGGCTGCTGCTGTCATCTGTGCCCCATCTGTGACCACTGCCGCCAGATAC 611
 Qy 153 LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGlu 172
 Db 612 CTGTCATCAATGATGTGTCTGTCTGGGCCCTGTCCCTGCTGCCAGATATCTGGAG 671
 Qy 173 TrpMetPheCysAbpPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
 Db 672 TGGATGTTCTGTGACCTTCGTTTAGTGCGTCGATTCGTGGTGTGAACCTCAGAT 731
 Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuVal 212
 Db 732 TTCATTACAAACGGGCGCTGGGTTTATTATGTTGTTCTCTGTGGGTCAGCTGGTC 791
 Qy 213 LeuLeuValAlaArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThr 232
 Db 792 CTGCTGTCAGGANTCTCTGTGATCCCGAAGATGCGCGACCAAGCTGTAGTACC 851
 Qy 233 IleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAla 252
 Db 852 ATCCTCTCACAGCTGCTGCTCTCTCTGTCGCTGCGCTTGGCATTCAGTGGGCC 911
 Qy 253 LeuPheSerArgIleHisLeuAspTrpValLeuPheCysHisValHisLeuValSer 272
 Db 912 CTGTTTCCAGATCCACCTGGATTTGAAGTCTTATTGTTCATGTGCATCTAGTTTCC 971
 Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValAlGlySer 292
 Db 972 ATTTCCTGTCCTCTCTTAACAGACAGTCCCAACCCCATTTACTTCTTGTGGGCTCC 1031
 Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuValLeuGlnArgAlaLeuGlnAsp 312
 Db 1032 TTTAGGCAAGCGTCAAAATAGGCAAGACCTGAAGCTGTTCTCCAGAGGGCTTCAGAGAC 1091
 Qy 313 ThrProGluValAlaArgGluGlyGlyIleTyrLeuProGlnGluThrLeuGluLeuSerGly 332
 Db 1092 ACGCCTAGAGTGAAGTGAAGTGAGGGTGGCTTCTCAGAGAAACCTGAGACTGTCCGGA 1151

```

QY      333 SerArgLeuGluGln 337
DB      1152 AGCAGATTGAGACG 1166

RESULT 4
US-10-183-116-30
; Sequence 30, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE, 4C1CPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (332)...(1297)
US-10-183-116-30

Alignment Scores:
Pred. No.: 1.7e-171 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 14 Gaps: 0

US-09-867-570-2 (1-337) x US-10-183-116-30 (1-1400)
QY      6 SerTPValIleArgLeuGlyPheLeuSerMetApSerThrIleProValIleGlyThr 25
DB      302 TTCAGGGTCACAGACTGGGGTTTCGAGCATGATTCACCACTCCAGCTTGGGTACA 361
QY      26 GluLeuThrProIleAsnGlyArgGluThrProCysEtyrIysGlnThrLeuSerPhe 45
DB      362 GAACGACACCAATCAACGACGACGACGACCTCTTGCACAAAGACCCGAGGCTTC 421
QY      46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValIleLeuTyr 65
DB      422 ACGGGGCTGACGTCATCGTTTCCTTTCGCGCTGACAGGAACGCGGTGGCTCTGG 481
QY      66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
DB      482 CTCCTGGGCTGCGCATGCGCAGAAACGCTGCTCAATTCACCTCAACCTGGTCGCG 541
QY      86 AlaAppPheLeuPheLeuSerGlyIleIleCysSerProLeuArgLeuIleAsnIle 105
DB      542 GCCGACTTCTCTTCTTTCGCGGCGCACATTAATGTTCCGCTTACCCCTCAATATC 601
QY      106 ArgHisProIleSerIysIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
DB      602 CGCCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCTTATATAGGCTCA 661

```

```

QY      126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTyrProIleTyr 145
DB      662 AGCATGCTAGCGCATTCAGCACCGAGCGCTCTGTCATCTTGGCCCATCTGGTAC 721
QY      146 HisCysArgArgProArgTyrLeuSerSerValMetCysValIleLeuTyrAlaLeuSer 165
DB      722 CACTGCGCGCGCCAGATACCTGTCATCGGTCATGTCGTCTCTCGGGCCCTGTC 781
QY      166 LeuLeuArgSerIleLeuGluTyrMetPheCysAppPheLeuPheSerGlyAlaAspSer 185
DB      782 CTGCTGCGAGATCTCTGAGTGAATGTTCTGTGACTTCTGTTTAAAGTGCTGATTC 841
QY      186 ValTyrCysGluThrSerAppPheIleThrIleAlaTyrPheValPheLeuCysValVal 205
DB      842 GTTGGTGTGAACGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 901
QY      206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
DB      902 CTCCTGGGTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
QY      226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
DB      962 CTGACCAGGCTGTACGTGACATCTCTCCACAGTGTGCTGCTCTCTCTGTCGCTG 1021
QY      246 ProPheGlyIleGluTyrAlaLeuPheSerArgIleHisLeuAspTyrPheValLeuPhe 265
DB      1022 CCGTTTGGCATTCAGTGGCCCTGTTTTCAGAGATCCACTGATTTGAAATCTTATTT 1081
QY      266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
DB      1082 TGTGATGTCATCTAGTTTCATTTCCATTTCTGTCGCTCTTAAACAGAGGCCAACCCATC 1141
QY      286 IleTyrPhePheValGlySerPheArgGlnArgIleAsnArgGlnAsnLeuVal 305
DB      1142 ATTACTTCTTCTGTCGCTCTCTTTCAGGACGCTCAAAATAGGCAACCTGAGT 1201
QY      306 LeuGlnArgAlaLeuGlnAspThrProGluValIaAspGluGlyGlyTyrLeuProGln 325
DB      1202 CTCGAGAGGCTCTGACAGACAGCTGAGTGTGATGAGTGAAGGTGGCTTCTCAG 1261
QY      326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
DB      1262 GAAACCTGAGAGCTGTCGGGAAACGATTTGAGCAG 1297

RESULT 5
US-10-225-567A-673
; Sequence 673, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biociences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 673
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-673

Alignment Scores:
Pred. No.: 1.7e-171 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2

```



```

Db      181  GTCTCATCTACATCTGACCTGGTCGGCGGCGAATTCTCTCTCTAGCGGCACATT 240
Qy      96  TLeCySerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241  ATATGTTGCGCGTTACGCTCATCAATATCCCGCATCCCATCTCCAAAATCTCTAGTCTCT 300
Qy      116  ValMetThrPheProTyPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301  GTGATGACCTTTCTCTTATTAAGCCTTAAGCACTGAGCGCCATCGACACGAGCCG 360
Qy      136  CysLeuSerIleLeuTrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSer 155
Db      361  TGCCGTGTCATCTGTGGCCCATCTGGTACACTGGCGCGCCGCGCCAGATACCTGTCATCG 420
Qy      156  ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      421  GTCATGTGTGTCCTGCTCTGGGCGCCCTGTCCCTGCGCGAGATACCTGAGAGGAGTTC 480
Qy      176  CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThr 195
Db      481  TGTGATCTTCTGTGTAGTGTGCTGATGTTGTGTGTGAGAAAGCTCAAGTTTCATTACA 540
Qy      196  IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
Db      541  ATCGGTGCGCTGTTTCTTTTATGTGTGTCTCTGTGTGTCCAGCTGTGCTGCTGCTGTC 600
Qy      216  ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db      601  AGGATTTCTCTGTGATGCCGGAAGATGCGCTGACGAGCTGTACCTGACCATCTCTCTC 660
Qy      236  ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSer 255
Db      661  ACAGGTGCTGTCTCTCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy      256  ArgIleHisLeuAspTrpIleValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db      721  AGGATCCACCTGTGATGGAAAGCTTATTTGTGTCATGTGATGATTTTCATTTCCTG 780
Qy      276  SerAlaLeuAspSerSerIleAsnProIleIleTyrPhePheValIleSerPheArgGlu 295
Db      781  TCCGCTCTTAAAGAGTGCACCAACCCCATCATTTACTTCTCTGTCGCTCTCTTAAAGCAG 840
Qy      296  ArgGluAsnArgGluAsnLeuIleValLeuGluArgAlaLeuGluIleAspTrpProGlu 315
Db      841  CGTCAAAATAGGCAAGAACTGAAAGCTGTCTTCCAGAGGCTTGCAGAGACAGCCTGAG 900
Qy      316  ValAspGluGlyGlyIleTrpLeuProGluIleThrLeuGluLeuSerGlySerArgLeu 335
Db      901  GTGATGAAAGTGTGAGGTGCTTCTCTCAGAAACCTGGAAGCTGTGCGGAAGCAGATTG 960
Qy      336  GluGlu 337
Db      961  GAGCAG 966

```

RESULT 7

```

US-09-995-225-19
; Sequence 19, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Friede, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human C
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13

```

```

; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent version 3.1
; SEQ ID NO 19
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: NO. US20030139588A9el Sequence
US-09-995-225-19

Alignment Scores:
Pred. No.: 2,99e-168 Length: 969
Score: 1688.00 Matches: 322
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.75% Indels: 0
DB: 10 Gaps: 0

US-09-867-570-2 (1-337) x US-09-995-225-19 (1-969)
Qy      16  MetAspSerThrIleProValLeuGlyThrGluLeuTrpProIleAsnGlyArgGluGlu 35
Db      1  ATGGATTCAACATCCCACTTGGGTGACAGACTGACCAATCAAGACGTAAGAG 60
Qy      36  ThrProCysTyrIleGluIleThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db      61  ACTCTTGCTACAGACAGACCTGAGCTTCAAGGGGCTGACCTGCATGCTTCCCTTATC 120
Qy      56  AlaLeuThrGlyAsnAlaValValLeuTrpLeuLeuGlyCysArgMetArgArgAsnAla 75
Db      121  GGGCTGACAGAAAGCGGGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy      76  ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db      181  GTCTCATCTACATCTCTCAACCTGTGCGGCGGCGCACTTCTCTTAAAGCGGCACATT 240
Qy      96  TLeCySerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241  ATATGTTGCGCGTTACGCTCATCAATATCCCGCATCCCATCTCCAAAATCTCTAGTCTCT 300
Qy      116  ValMetThrPheProTyPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301  GTGATGACCTTTCTCTTATTAAGCCTTAAGCACTGAGCGCCATCGACACGAGCCG 360
Qy      136  CysLeuSerIleLeuTrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSer 155
Db      361  TGCCGTGTCATCTGTGGCCCATCTGGTACACTGGCGCGCCGCGCCAGATACCTGTCATCG 420
Qy      156  ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      421  GTCATGTGTGCTCTGCTCTGGGCGCCCTGTCCCTGCTGCGAGATACCTGAGAGGAGTTC 480
Qy      176  CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThr 195

```

Db	481	TGTGACTTCCTGTTAGTGTGCTCTAATCTGTTGGTGTGAAACGTACAGTTTCATTACA	540
QY	136	ILLeaIaTRPLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal	215
Db	541	ATCGCGTGGCTGGTTTTTTTATATGTGTGGTCTCTGTGGGTCCAGCCTGGCTGCTGCTC	600
QY	216	ArgIleLeuCysGlySerArgIysMetProLeuThrArgLeuIyrValThrilLeuLeu	235
Db	601	AGGATTTCTGTGTGATCCCGAAGATGCGCGCTGACCAAGGTGTACGTGACATCTCTCTC	660
QY	236	ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGIInTPalaleuPheSer	255
Db	661	ACAGTGGTGGTCTTCTCTCTCTGTGGCCGTGCCCTTTGGCATTCAGTGGGGCCGTGTTCC	720
QY	256	ArgIleHisLeuAspTrpIysValLeuPheCysHisValHisLeuValSerIlePheLeu	275
Db	721	AGGATCCACCTGGATTTGAAAGCTTATTTTGTGCATGTGCATCTAATTTTCATTCTCTG	780
QY	276	SerAlaLeuAsnSerSerIalaAsnProIleIleTyrPhePheValIleSerPheAspGln	295
Db	781	TTCGGCTCTTAACAGCAGTGTCCAAACCCATCATTTACTTCTTCGTGGGCTCTTTTAGGCAG	840
QY	296	ArgGlnAsnArgGlnAsnLeuIysLeuValLeuGlnArgAlaLeuGlnAspTrpProGlu	315
Db	841	CCTCAAAATAGGCACAGACTGAAAGCTGGTCTTCCAGAGGGCTTGCAGAGACAGCCTGAG	900
QY	316	ValAspGluGlyGlyGlyTyrTrpLeuProGlnIuThrLeuGluLeuSerGlySerArgLeu	335
Db	901	GTCGATGAAGGTGGAGGGTGGCTTCTCTCAGAAACCTGGAGCTGTCTGGAGACAGATTG	960
QY	336	GIuGln 337	
Db	961	GAGCAG 966	

RESULT 8

```

US-10-401-397A-1
: Sequence 1, Application US/10401397A
: Publication No. US20030212001A1
: GENERAL INFORMATION:
: APPLICANT: Peri, Krishna G.
: APPLICANT: Moffett, Serge
: APPLICANT: Abtran, Daniel
: TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION OF ELEVATED
: TITLE OF INVENTION: INTRAOCULAR PRESSURE AND RELATED CONDITIONS
: FILE REFERENCE: 4518/1M674US1
: CURRENT APPLICATION NUMBER: US/10/401,397A
: CURRENT FILING DATE: 2003-03-27
: PRIOR APPLICATION NUMBER: US 60/367,513
: PRIOR FILING DATE: 2002-03-27
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 969
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(969)
: OTHER INFORMATION: coding sequence for polypeptide
US-10-401-397A-1

```

Alignment Scores:

Pred. No.:	1.0e-167	length:	9650
Score:	1683.00	Matches:	321
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	95.46%	Indels:	0
DB:	15	Gaps:	0

US-09-867-570-2 (1-337) x US-10-401-397A-1 (1-969)

QY 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGlu 35

Db	1	ATGATGTTCAACCATCTCCAGTCTTGSGTACAGAACTGACCAATCAACGAGCGTAGGAG	60
Qy	36	ThrProCysTyrIYVGLnThrlensePherThrgIYleuthCysIIeValSerLeuVal	55
Db	61	ACTCCTTGCTCAAGCAAGCAACCCCTGAGCTTCAACGGGGCTGACCTGATCCCTTCCCTTGTC	120
Qy	56	AlaLeuThrgIYAsnAlIValIValLeuThrlPheLeuGlyCYsArgmCcrArgAsnAla	75
Db	121	GGCGTGAAGAGAGACGGCGTGTGTCTGCGCTCTCGGGCTGGCGCATGGCAGAACGCT	180
Qy	76	ValSerIleYrIIeLeuAsnLeuValAlaAspPheLeuPheLeuSerGlyHisIle	95
Db	181	GTCCTCACTCAATCTCAACCTGGTCCGGCGGCGCACTTCTCTTCCCTTGCGGCGCAATY	240
Qy	96	IleCysSerProLeuArgLeuIleAsnIIeAspHisProIleSerIYsIIeLeuSerPro	115
Db	241	ATATGTTGCGGTACGGCTCAATCAATACGGCAATCCCATCTCCAAATTCCTCACTCT	300
Qy	116	ValmetThrPheProYrPheIIeGlyLeuSermetLeuSerAlaIleSerThrgIuArg	135
Db	301	GTGATGACCTTTCCTTATATAGACCTTAAGCATGCGAGCGCATGACACCGAGCGC	360
Qy	136	CysLeuSerIIeLeuTrpProIIeTrpYrHisCYsArgArgProArgYrIleuSerSer	155
Db	361	TGCCTGTCCATCTGTGGCCCATCGTGAACCATCCCGCGGCCAGATACCTGTATCG	420
Qy	156	ValmetCysValIleLeuThrlAlaLeuSerLeuLeuArgSerIIeLeuGluTrpmetPhe	175
Db	421	GTCAATGTGTCTCTGCTCTGGGCGCTGTCCCTGCTGCGAGATCTCTGAGTGTGATCTTC	480
Qy	176	CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAspPheIIeThr	195
Db	481	TGTGACTTCCGTGTAGTGTGTGCTATCTGTGTGTGAACGTCAGATTCATTACATA	540
Qy	196	IleIleIleThrlPheLeuValPheLeuCYsValIleLeuCYsGlySerSerIleuValIleuVal	215
Db	541	ATCCGATGGCGTGTATTATTATGTGTGTCTCTGTGGTCCAGCGCTGTCTCGCTGTTC	600
Qy	216	ArgIIeLeuCYsGlySerArgIYsMetProLeuThrArgLeuYrValThrlIleLeuLeu	235
Db	601	AGGATTCCTCTGTGATCCCGGAAGATGCGGTGACAGCGCTGTATCGTGAACATCTCTTC	660
Qy	236	ThrValIleuValPheLeuLeuCYsGlyLeuProPheGlyIIeGluThrlAlaLeuPheSer	255
Db	661	ACAGTGCAGTCTCTCTCTCTGTGGCCGTGCCCTTGGCATTCAGTGGGCCCTGTCTTCC	720
Qy	256	ArgIIeHisLeuAspTrpIYsValIleuPheCYsHisValHisLeuValSerIIePheLeu	275
Db	721	AGGATCCACCTGGATGGAAAGTCTTATTTGTCTATGTGCATCTAGTCTTCATTTCCCTG	780
Qy	276	SerAlaLeuAsnSerSerAlaAsnProIleIleYrPhePheValGlySerPheArgIn	295
Db	781	TCCCTCTTTACAGACAGTGCACACCCCATCTTACTTCTTGTGTGGCTCTTATAGGACAG	840
Qy	296	ArgGlnAsnArgGlnAsnLeuIYsLeuValIleuGlnArgAlaLeuGlnAspThrProGlu	315
Db	841	CGTCAAAATAGCGCAACCTGAAGTGTGTCTTCCAGAGGGCTCTGAGGACACGCGCTGAG	900
Qy	316	ValAspGluGlyGlyIYTrpLeuProGluGluThrlLeuGluLeuSerGlySerArgLeu	335
Db	901	GTGATGTAAGGTGAGGGGTGCTTCTCAGAGAAACCTGAGAACTGTGCGGAGACATTTG	960
Qy	336	GluGln 337	
Db	961	GAGCAG 966	

RESULT 9

```

; Sequence 1, Application US/10391074
; Publication No. US20040038345A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael

```

APPLICANT: Mattler, Frank
 TITLE OF INVENTION: No. US20040038345A1el Human Seven-Transmembrane Receptors
 FILE REFERENCE: 7705.0008-00-000
 CURRENT APPLICATION NUMBER: US/10/391,074
 CURRENT FILING DATE: 2003-03-17
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 969
 TYPE: DNA
 ORGANISM: Homo sapien
 US-10-391-074-1

Alignment Scores:

Pred. No.:	1,316-165	Length:	969
Score:	1663.00	Matches:	320
Percent Similarity:	99.38%	Conservative:	0
Best Local Similarity:	99.38%	Mismatches:	2
Query Match:	94.33%	Indels:	0
DB:	16	Gaps:	0

US-09-867-570-2 (1-337) x US-10-391-074-1 (1-969)

```

Qy 16 MetaSpSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGln 35
Db 1 ATGGATTCAACATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGAGTGAGAG 60
Qy 36 ThrProCysTyrIleGlnThrIleuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db 61 ACTCTTCTCTACACAGACAGACCTGAGCTTACCGGGCTACGCGATCTCTTCTTGC 120
Qy 56 AlaIleuThrGlyAsnAlaValIleuThrPheLeuGlyCysArgMetArgAsnAla 75
Db 121 GCGCTGACAGGAAAGCGGTGTGCTCTTCTGCTCTGAGCTGCGGCTGCGGAGGAGGAG 180
Qy 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyIle 95
Db 181 GTCTCATCTACATCTCTCAACCTGCGGCGGCGGCACTTCTCTTCTTACGCGCACATT 240
Qy 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db 241 ATACCTTGGCCCTTACGCTCATCATATCCGCTATCCGCAATCCCAATCTCACTGCT 300
Qy 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db 301 GTGATGACCTTCTCTACTTATAGGCTTACAGATGCTAGGCCCATCAGCACCGAGCGC 360
Qy 136 CysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyrIleuSerSer 155
Db 361 TGGCTGTCATCTGTGGCCCATCTGTGACCACTGCGCGCGCCGACAGATCACTGTCATCG 420
Qy 156 ValMetCysValLeuLeuThrPalaLeuSerIleLeuArgSerIleLeuGluThrPhe 175
Db 421 GTGATGTGTGCTCTGCTGTGGCGCTGCTGCTGCTGCGAGTATCTTGGAGTGTTC 480
Qy 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db 481 TGTGACTTCTCTTATAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 196 IleAlaThrPheLeuValPheLeuCysValIleLeuCysGlySerSerIleValIleLeuVal 215
Db 541 ATCGGTGGCTGCTGCTTATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuVal 235
Db 601 AGGATCTCTGTGGATCCCGAGAGATGCGCTGACCAAGGCTGTACCGATCACTCTCTC 660
Qy 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrPalaLeuPheSer 255
Db 661 ACGAGTGTGTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy 256 ArgIleLeuValPheLeuSerValIleuPheCysHisValHisLeuValSerIlePheLeu 275
  
```

```

Db 721 AGGATCAACCTGATGGAAAGTCTTATTTGTACATGTCATCTAGTTCCATTTCTCTG 780
Qy 276 SerAlaLeuAspSerThrIleAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db 781 TCCGCTCTTAACAGAGTGCACCCCACTTATCTTCTGTGGGCTCTTATAGGAG 840
Qy 296 ArgGlnAsnArgGlnAsnLeuLeuValIleuGlnArgAlaLeuGlnAspThrProGlu 315
Db 841 CGTCAAAATAGGCACAACTGAGAGCTGTTCTCCAGAGGGCTCTGCAGACACGCTGAG 900
Qy 316 ValAspGluGlyGlyTyrPheuProGlnGluThrIleGluLeuSerGlySerArgLeu 335
Db 901 GTGGATGAAGTGTGAGGCGAGCTTCTCAGGAAACCTGAGCTGTCTGGAGAGAGATTG 960
Qy 336 GluGln 337
Db 961 GAGCAG 966
  
```

RESULT 10

US-10-237-467-9
 Sequence 9, Application US/10237467
 Publication No. US20030186324A1

GENERAL INFORMATION:

APPLICANT: Liao, JiaYu
 APPLICANT: Gray, Nathanael S.
 APPLICANT: Caldwell, Jeremy C.
 APPLICANT: Schultz, Peter G.

APPLICANT: IRM LLC

TITLE OF INVENTION: Sensory Neuron Receptors

FILE REFERENCE: 021288-001300US

CURRENT APPLICATION NUMBER: US/10/237,467

CURRENT FILING DATE: 2003-01-14

PRIOR APPLICATION NUMBER: US 60/317,879

PRIOR FILING DATE: 2001-09-07

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 909

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE: OTHER INFORMATION: dorsal root ganglia G-protein coupled receptor (GPCR)

OTHER INFORMATION: 7 (DRG7) (NT009307)

US-10-237-467-9

Alignment Scores:

Pred. No.:	1,886-154	Length:	909
Score:	1557.00	Matches:	301
Percent Similarity:	93.79%	Conservative:	1
Best Local Similarity:	93.48%	Mismatches:	0
Query Match:	88.32%	Indels:	20
DB:	15	Gaps:	1

US-09-867-570-2 (1-337) x US-10-237-467-9 (1-909)

```

Qy 16 MetaSpSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGln 35
Db 1 ATGGATTCAACATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGAGTGAGAG 60
Qy 36 ThrProCysTyrIleGlnThrIleuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db 60 -----
Qy 56 AlaIleuThrGlyAsnAlaValIleuThrPheLeuGlyCysArgMetArgAsnAla 75
Db 61 GCGCTACAGGAAACGCGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyIle 95
Db 121 GTTCTCATCTACATCTCTCACTGTGTGCGGCGGCACTTCTTCTTACGCGCACATT 180
Qy 180
  
```



```

Db      181 ATATGTTGGCGCTACGCTCATCATATCCGCCATCCCATCTCCAAATCTCATGCTCT 240
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      241 GTGATGACCTTCCCTCACTTATAGCCCTTAAGCATGTCAGACGCCCATCAGCAGGAGGC 300
Qy      136 CysLeuSerIleLeuTrpProIleTrpTyrHisGlyArgAlaProAlaGlyTyrLeuSerSer 155
Db      301 TGCCTGTCACTCCGTGGCCCATCTGTACACACGCGCCGCCAGATACCTTCATCA 360
Qy      156 ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      361 GTATGTGTGTCCGCTCTGGGCTGTGCTGCTGCTGCGGATATCTCGAGTGCATGCTTC 420
Qy      176 CysAspPheLeuPheSerGlyValAlaAspSerValTrpCysGluTrpSerAspPheIleThr 195
Db      421 TGTCACTTCCCTGTTAGTGGTGAATATCTGTTGTGTGAACGTCAAGATTTTCATTACA 480
Qy      196 IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
Db      481 ATCCGCTGTGTGTGATCCCGAAGATGCCGCTGACAGGCTGTACGTGACCATCTCTC 540
Qy      216 ArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db      541 AGGATTTCTGTGTGATCCCGAAGATGCCGCTGACAGGCTGTACGTGACCATCTCTC 600
Qy      236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSer 255
Db      601 ACAGTGTGTGCTTCTCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      256 ArgIleIleLeuAspTrpPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db      661 AGGATTCACCTGAGTGTGAAGTCTTATTTGTGATGATGATGATGATGATGATGATGATGAT 720
Qy      276 SerAlaLeuAspSerSerAlaAspProIleIleTyrPhePheValGlySerPheArgGln 295
Db      721 TCCGCTTTAAACGACATGCGAACCCCATCTTACTTCTTCGTGGGCTCTTTAGGCG 780
Qy      296 ArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAspThrProGlu 315
Db      781 CGTCAAAATAGGCGAAGACCTGAACTGATGCTCTCCAGAGGCTCTGACGACACCGCTGAG 840
Qy      316 ValAspGlnGlyGlyTyrTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
Db      841 GTGATGTAAAGTGTGAGGTGTGCTTCTCAGGAAACCTGTGAGCTGTGGGAAAGCATGTG 900
Qy      336 GluGln 337
Db      901 GAGCAG 906

```

```

RESULT 11
US-10-017-161-1055
; Sequence 1055, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEYANATHAN, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017, 161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1055
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source

```

```

; LOCATION: (1)..(1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1169)
US-10-017-161-1055

Alignment Scores:
Pred. No.: 2,73e-139 Length: 1369
Score: 1416.00 Matches: 277
Percent Similarity: 88.25% Conservative: 16
Best Local Similarity: 83.43% Mismatches: 39
Query Match: 80.32% Indels: 0
DB: 15 Gaps: 0

US-09-867-570-2 (1-337) x US-10-017-161-1055 (1-1369)
Qy      6 SerTrpValIleArgGlyLeuPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db      171 TCCAGGCTCACCACTAAGGTTTCTAGCATGATCCACCATCTCAACCTTGACACA 230
Qy      26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrGlyGlnThrLeuSerPhe 45
Db      231 GAACGTACACCAATCAACGGAACCTGAGAGACTCTTTGCTACAGACAGACCTTGAGCTTC 290
Qy      46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValIleuTrp 65
Db      291 ACGGTGTGACGTGCATCTGTTCCCTGTGCGGCTGACAGAAACGAGTGTCTCTG 350
Qy      66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db      351 CTCTGGGCTGCGCCATGCGGAGAGAGCCCTTCTTCATCTACATCTTCAACTTGCGCGCA 410
Qy      86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db      411 GCAACCTTCTCTCTCTCAGAGGCGGCTTATATATTCCTGTTAAGCTTCATCAGATC 470
Qy      106 ArgHisProIleSerIleValLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db      471 CCCCATACCACTCTTAAATCTCTATCTGTATGATGATGATGATGATGATGATGATGATGAT 530
Qy      126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTyr 145
Db      531 AGCTTTGAGTGTGCGGAGGACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
Qy      146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer 165
Db      591 CGCTGCAACCGCCCAACACACACCTGTACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
Qy      166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Db      651 CTGCTGCGGAGACATCTCGAGGTGATGTATGTGCTCTCTGTTCACTGTGCTGCTGATCTT 710
Qy      186 ValTrpCysGluTrpSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Db      711 GCTTGTGTCAACATGATGATTCATCAGAGTCCGCGGCTGATTTTATGTGTGTT 770
Qy      206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgGlyMetPro 225
Db      771 CTCTGTGTGTGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
Qy      226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
Db      831 CTGACCAAGGCTGTACGACCATCTCTCAGAGTACGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 890
Qy      246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpValLeuPhe 265
Db      891 CCTTTGGATTCAGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 950
Qy      266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAspSerAlaAspProIle 285
Db      951 TGTCAATTCATGATTTCTATTTCTCTGCTGCTGCTTAAACGACGACGACGACGACGACGAC 1010
Qy      286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuVal 305

```

```

Db      1011 ATTACTCTTCCTGCTGGCTCTTTCAGGACGCTCAAAATAGCAGAACTGAAAGCTGTGTT 1070
Qy      306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db      1071 CTCAGAGGGCTCTCTCAGGACCGCTGTGAGTGAGTGAAGTGAAGGAGGAGCTTCTCTGAG 1130
Qy      326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db      1131 GAAATCTGTGAGCTGTGGGAAAGCAGATTGAGACAG 1166

```

RESULT 12

```

US-10-292-798-897
; Sequence 897, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AHIRATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 897
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1) .. (1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201) .. (1169)
; US-10-292-798-897

```

Alignment Scores:

```

Pred. No.:      2,73e-139      Length:      1369
Score:          1416.00      Matches:      277
Percent Similarity: 88.25%      Conservative: 16
Best Local Similarity: 83.43%      Mismatches: 39
Query Match:      80.32%      Indels:      0
Db:              15          Gaps:      0

```

US-09-867-570-2 (1-337) x US-10-292-798-897 (1-1369)

```

Qy      6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db      171 TCCAGGGTCACCAAGACTAGAGTTTCTGAGATGATCAACATCTCAACCTTGAGACACA 230
Qy      26 GluLeuThrProIleAsnGlyArgGluGluThrProCysGlyTyrIleGlnThrLeuSerPhe 45
Db      231 GAATGACACCAATCAACGGAACCTGAGAGACTCTTGTGCAAGAGACCTTGAGCGCTC 290
Qy      46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyValAsnAlaValIleLeuTrp 65
Db      291 ACGGTCGTCACGTCATCTGTTCCCTTTCGCGCTACAGAGAAACGAGTTGTGCTCTG 350
Qy      66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db      351 CTCCTGGGCTCCGATGCGAGCAAGACGCTTCTCCATCTACATCTCAACTTGCGCGCA 410
Qy      86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db      411 GCAAGCTTCCTCTCTCTCAGCGCGCGCTTATATATATCTCTCTGTTAAGCTTCATCAGTATC 470

```

```

Qy      106 ArgHisProIleSerIleIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db      471 CCCCATACCATCTCTAAATCTCTATCTCTGATGATGATGTTTCTTACTTTCAGAGCCTG 530
Qy      126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTyr 145
Db      531 AGCTTCTGAGCGCGTGAACACGAGCGCTCCTGTCGCTCTGCGCCATCTGATGATC 590
Qy      146 HisCysArgArgProArgTyrIleLeuSerSerValMetCysValIleLeuTrpAlaLeuSer 165
Db      591 CGCTGCCACCGCCCAACACCTGTCAGCGGTGTGTGTCTCTGCGCCTGCTGCC 650
Qy      166 LeuLeuArgSerIleLeuGluThrPheMetPheCysAspPheLeuPheSerGlyValAspSer 185
Db      651 CTGCTCGGAGCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 710
Qy      186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Db      711 GCTTGGTGTCAAAACATCAGATTTCATCAGCGGTGCTGATTTTATGTGTGTT 770
Qy      206 LeuCysGlySerSerLeuValLeuLeuValAlaGlySerArgIleSerArgIleSerPro 225
Db      771 CTCTGGGTTCAGCGCTGCTCTGATCAGATTCCTGTGATCCCGGAAGATTCG 830
Qy      226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
Db      831 CTGACCAAGCTGTACGTGACATCTCTGCTCAAGATCTGTTCTCTCTCTGTCGCTG 890
Qy      246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIleValLeuPhe 265
Db      891 CCTTTGGCATTCAGTTTCTCTATTTTATGATCCACGTGACAGGGAAGTCTTATTT 950
Qy      266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
Db      951 TGTCAATGTCATCTATGTTCTATTTCTCTGCTCTTAAAGAGATGCAACCCATTC 1010
Qy      286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuLeuVal 305
Db      1011 ATTACTTCTTCTGTGGCTCTTTAGGACGCTCAAAATAGCAGAACTGAAAGCTGTGTT 1070
Qy      306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db      1071 CTCAGAGGGCTCTCTCAGGACCGCTGTGAGTGAGTGAAGTGAAGGAGGAGCTTCTCTGAG 1130
Qy      326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db      1131 GAAATCTGTGAGCTGTGGGAAAGCAGATTGAGACAG 1166

```

RESULT 13

```

US-10-183-116-15
; Sequence 15, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melyun
; APPLICANT: Han, Sang-lyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE, 4C1CPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109

```

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15
 LENGTH: 2040
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (328)...(1293)
 US-10-183-116-15

Alignment Scores:

Pred. No.:	4,98e-139	Length:	2040
Score:	1416.00	Matches:	277
Percent Similarity:	88.25%	Conservative:	16
Best Local Similarity:	83.43%	Mismatches:	39
Query Match:	80.32%	Indels:	0
	14	Gaps:	0

US-09-867-570-2 (1-337) x US-10-183-116-15 (1-2040)

```

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValIleuGlyThr 25
Db 298 TCCAGGAGCTACACAGCTAGAGGTTCTGAGCATGATCCAACTCTCAACCTTGAGACCA 357

Qy 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrIleGlnThrLeuSerPhe 45
Db 358 GAATGACACCAACCAACGAGAACTGAGAGACTCTTTGCTACAGACAGACCTTGAGCCTC 417

Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValIleuTrp 65
Db 418 ACCGTCGTGACGTGATCTGTTCCCTTGTGCGGCTGACAGAAACGAGTTGCTCTGG 477

Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 478 CTCCTGGGCTGCCCGACATGCGAGAGAGCCTTCTCCATCTTCATCTCAACTTGAGCGCA 537

Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db 538 GCAAGACTTCTCTTCTCAGCGCGCGCTTATATATTCCTGTTAAGCTTCATCAGATAC 597

Qy 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db 598 CCCCATACCATCTTAATCTCTATCTCTGTATGATGTTTCTTCTTGGACGGCTG 657

Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTyr 145
Db 658 AGCTTCTGAGTGCGTGAGACCGAGGCGCTGTCGCCGCTGTCGAGCCATCTGGATC 717

Qy 146 HisGlyArgArgProArgTyrLeuSerSerValMetCysValIleLeuTrpAlaLeuSer 165
Db 718 CGCTGCCACCGCCACACACACTGTCAAGCGGTGTGTGTCCTGCTGCGGCCCTGCTCC 777

Qy 166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Db 778 CTGCTGGGAGACATCTCTGAGATGATGTATGTGCTTCTCTTCAAGTGGTGCATTTCT 837

Qy 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Db 838 GCTTGAGTCAAAATCAGATTCATCAAGTCGCGGCTGATTTTATGTGTGTT 897

Qy 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
Db 898 CTCTGTGGGTCCACCGTCTCTGCTGATCAGATTCCTGTGTGATCCCGAATATACCG 957

Qy 226 LeuThrArgLeuValThrIleLeuLeuThrValIleuValPheLeuLeuCysGlyLeu 245
Db 958 CTGACGACGCTGTCGACATCTCTCTCACTACGATCTCTCTCTCTGCGCTG 1017

Qy 246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIleValLeuPhe 265
Db 1018 CCCCTTGGCATTCAGTTTCTTATTTTATGATCCACGCGACAGGAGATCTTATTT 1077

Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerAlaAsnProIle 285

```

```

Db 1078 TGTCAATTCATCTAGATTTCTATTCTTCTCCCTTTAAACAGACATGCCAACCCATC 1137

Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgIleAsnLeuValLeuVal 305
Db 1138 ATTACTTCTTCCGTGGGCTCTTTAGCAGCGTCAAAATAGGCAAGACCTGAGCTGGTT 1197

Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db 1198 CTCGAGGCGCTCTGACGAGACGCTCTGAGTGATAGATGAGGAGGAGCTTCTGAG 1257

Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db 1258 GAATCTCTGAGCTGTCTCGAGAGCATGTGAGACG 1293

RESULT 14
US-10-017-161-1599
: Sequence 1599, Application US/10017161
: Publication No. US20030143668A1
: GENERAL INFORMATION:
: APPLICANT: SUMA, MAKIKO
: APPLICANT: ASAI, KIYOSHI
: APPLICANT: AKIYAMA, YUTAKA
: APPLICANT: ABURATANI, HIROYUKI
: TITLE OR INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
: FILE REFERENCE: 084335/0152
: CURRENT FILING DATE: 2002-12-18
: PRIOR APPLICATION NUMBER: JP 2001/246789
: PRIOR FILING DATE: 2001-06-18
: NUMBER OF SEQ ID NOS: 2430
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1599
: LENGTH: 1370
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: source
: LOCATION: (1)..(1370)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (201)..(619)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (966)..(1170)
: US-10-017-161-1599

Alignment Scores:
Pred. No.: 7.3e-138 Length: 1370
Score: 1402.50 Matches: 293
Percent Similarity: 87.93% Conservative: 13
Best Local Similarity: 84.20% Mismatches: 29
Query Match: 79.55% Indels: 15
DB: 15 Gaps: 4

US-09-867-570-2 (1-337) x US-10-017-161-1599 (1-1370)
Qy 2 GluSerIleSerSerTrp-----ValIleArgLeuGly 12
Db 132 CAGAGATCAAAACAGCTGTGATCATCATCTGTTCTCTTCCAGGCTCATCAGCTGGGG 191

Qy 13 PheLeuSerMetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGly 32
Db 192 TTTCTGAGCATGATTCACCATCTCCAGTCTTGGGTACAGAACTGACCAATCAACGGA 251

Qy 33 ArgGluGluThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 252 CGTAGGAGACTCTTCTCAAGACAGACCTGAGCTTCAAGGAGCTGACGTGATGCTT 311

Qy 53 SerLeuValAlaLeuThrGlyAsnAlaValIleuTrpLeuLeuGlyCysArgMetArg 72
Db 312 TCCCTTCTCGCGCTGACAGGAAACGCGTTGTGCTCTCGGCTCGGCTCGCATGCGC 371

```

Oy	73	ArgAnnaIaValSerIleTyrIleLeuAsnLeuValAlaIaAspPheLeuPheLeuSer	92
Db	372	AGGAAGCGTGTCTCCATCTACATCTCTCAACCTGGTGGCGGCACTTCTCTCTTACG	431
Oy	93	GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle	112
Db	432	GAGCCACATTAATATGTTCGCCGTACCGCTCATCAATATCCGCCATCCATCCCAAAATC	491
Oy	113	LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer	132
Db	492	CTCAATCTCTGTGATACCTTCTTCCCTACTTTATAGGCTTAAGCATCTGAAGCCATCAGC	551
Oy	133	ThrGluArgCysLeuSerIleLeuTyrProIleTyrHisCysArgArgProArgTyr	152
Db	552	ACCGAGCGCTGCTGTCCATCTGAGCCCATCTGGTACACTC-CGCGGCCCGAGATAC	610
Oy	153	LeuSerSerValMetCysValLeuLeuTyrPalalaLeuSerIleLeuArgSerIleLeuGlu	172
Db	611	CTGTATCGGTATGTATGTCTGTCTGTGTGTCCTGTCTCCCTGTCTCGGAGATCATGAG	670
Oy	173	TyrMetPheCysAspPheLeuPheSerGly--AlaAspSerValTyrCysGluThrSer	191
Db	671	TGGATGATCTGTGACTACTCTGTGAGTGTGTCTGTCTAATGCGGGTGGTGTAAAGCACCA	730
Oy	192	AspPheIleThrIleAlaTyrLeuValPheLeuCysValValLeuCysGlySer--Ser	210
Db	731	GAC--ATTACAAACGCGGTGCTGTGTGTATATAGAGGGGTCTCGCGGGGGGTGCAGC	787
Oy	211	LeuValLeuLeuValArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeu-Ty	230
Db	788	GTCTGCTCTGTCTATAGATATGGCGGAGATATGCGGAAATGCCCCCTAGAGATGTTCCAG	847
Oy	230	TyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleG	250
Db	848	CGTGATCATCTCCCTCATATGATGTGTGTCTCTCCGCGGTGCTCGCTTTGG--ATTCA	905
Oy	250	nTyrPalalaLeuPheSerArgIleHisLeuAspTyrIleValLeuPheCysHisValHisIle	270
Db	906	GTGGGCGCTGTTTTCAGGATCCACCTGGAGTGAAGTCTTATTTGTGCATGTCACTCT	965
Oy	270	uValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheVa	290
Db	966	AGTTTCATTTTCTGTGCGCTTAAACGACGTGCACACCCATCATTTACTTCTTCTGT	1025
Oy	290	IGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuIleValLeuGlnArgAlaIle	310
Db	1026	GGGCTCCCTTTAGCAGCGCTCAAAATAGGAGAACTGAAAGCTGTCTCCAGAGGGCTCT	1085
Oy	310	uGlnAspThrProGluValAspGluGlyGlyTyrLeuProGlnGluThrLeuGluIle	330
Db	1086	GCAGACACGCTTGAAGTGGAGTGAAGGTGGAGGTCTTCTCAGAAACCTGTGAGACT	1145
Oy	330	uSerGlySerArgLeuGlnGln	337
Db	1146	GTCCGGAAAGACAGATTGGAGCAG	1167
RESULT 15			
US-10-072-012-171			
; Sequence 171, Application US/10072012			
; Publication No. US2004003493A1			
; GENERAL INFORMATION:			
; APPLICANT: Tchernev, Valizar			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zethusen, Bryan			
; APPLICANT: Patursajan, Meera			
; APPLICANT: Shimkets, Richard			
; APPLICANT: Li, Li			
; APPLICANT: Gangoli, Esna			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Anderson, David W.			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Miller, Charles E.			
; APPLICANT: Gerlach, Valerie			

```

APPPLICANT: Taupler Jr, Raymond J.
APPPLICANT: Gusev, Vladimir Y.
APPPLICANT: Coleman, Steven D.
APPPLICANT: Wolenc, Adam R.
APPPLICANT: Pena, Carol E. A
APPPLICANT: Furtak, Katarzyna
APPPLICANT: Grose, William M.
APPPLICANT: Alebrook II, John P.
APPPLICANT: Lepley, Denise M.
APPPLICANT: Rieger, Daniel K.
APPPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining prior ID Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 171
LENGTH: 997
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-171

Alignment Scores:
Pred. No.: 1,72e-137 Length: 597
Score: 1397.00 Matches: 272
Percent Similarity: 88.62% Conservative: 16
Best Local Similarity: 83.69% Mismatches: 37
Query Match: 79.24% Indels: 0
DB: 16 Gaps: 0

US-09-867-570-2 (1-337) x US-10-072-012-171 (1-997)
QY 13 PheuserSerMetAspSerThrIleProValIleuGIYThrGIUleuThrProIleasnGIY 32
DB 3 TTTCTGAGCATTGATTCACCAACATCTCAACTTGACACAGAACTGACCACTCAACGA 62
QY 33 ArgGIUleuThrProCYsTyrIysGIInThrLeuSerPheThrGIYleuThrCYsIleVal 52
DB 63 ACTGAGGAGACTCTTTGCTCAACAGCAGCACTTGACCTCAAGGCGTGCAGACCTGATGTT 122
QY 53 SerIeuValIaleuThrGIYasnIaValIalleuThrPleuLeuGIYCYsArgMetArg 72
DB 123 TCCCTTGTCCGCTGACAGGAACGACGTGTGCTCTGCTCTCCGCTGCCGATGCC 182
QY 73 ArgasnIaValSerIleTyrIleuLeuIeuValIaIaIaAspPheLeuPheLeuSer 92
DB 183 AGGAACGCTTCTTCATCTCAATCCCTCAACTTGCGCGCAGAGACCTCTCTCTCAGC 242
QY 93 GlyHisIleIleCYsSerProIeuArgIleuIleasnIleArgHisProIleSerIle 112
DB 243 GGGCGCTTAAATATTCCTCGTAAGCTTCATCAAGATATCCCCCAATCAATCTTAATTC 302

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 1, 2004, 21:59:08 / Search time 3989 Seconds

(without alignments)
3995.144 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
Sequence: 1 MESKSSVIVILGFLSHMSTI.....EGGGMPLQRTLELSGSLLEQ 337

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US09867570/runat.01122004.161122.6337/app.query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LCOPLC=0 -LCOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09867570.OCGN.1.1.5600.0/runat.01122004.161122.6337 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hvg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	2618	6	AX429465 Sequence
2	1727.5	98.0	8622	6	AX429467 Sequence
3	1722.5	97.7	1369	6	AX647081 Sequence
4	1722.5	97.7	91923	9	AC108007 Homo sapi

C	5	1722.5	97.7	172939	9	AC090099
	6	1721	97.6	1400	6	AX299158
	7	1721	97.6	1400	6	AX549388
	8	1721	97.6	1400	6	AY044215
	9	1700	96.4	1697	9	BC067292
	10	1688	95.7	969	6	E43450
	11	1688	95.7	969	6	E43450
	12	1683	95.5	969	6	E43451
	13	1683	95.5	969	6	AX923125
	14	1675	95.0	969	6	AF474987
	15	1642	93.1	969	6	AR477199
	16	1642	93.1	969	6	BD103571
	17	1606	91.1	969	6	AF474988
	18	1598	90.6	969	6	AR477200
	19	1598	90.6	969	6	BD103572
	20	1416	80.3	1369	6	AB065846
	21	1416	80.3	1369	6	AB065846
	22	1416	80.3	2040	6	CO815414
	23	1416	80.3	2040	6	AX299143
	24	1416	80.3	2040	9	AY042213
	25	1416	80.3	156839	9	AC107948
C	26	1416	80.3	153718	9	AC023078
C	27	1416	80.3	153718	2	AC139482
	28	1411	80.0	969	9	AF474989
	29	1405	79.7	969	6	AR477201
	30	1405	79.7	969	6	BD103573
	31	1402.5	79.6	192618	2	AC020568
	32	1399	79.4	912	6	AX657514
	33	1397	79.2	997	6	AX921831
	34	1390	78.8	1369	6	AX646849
	35	1388	78.7	1604	6	AX299160
	36	1388	78.7	1604	6	AX549403
	37	1388	78.7	1604	9	AY042216
	38	1387	78.7	969	6	E40003
	39	1387	78.4	969	6	AX148178
	40	1383	78.4	969	6	AX268489
	41	1383	78.4	969	6	AX375428
	42	1383	78.4	969	6	AX671113
	43	1383	78.4	969	6	BD144297
	44	1383	78.4	969	6	AB083628
	45	1383	78.4	1116	6	BD095749

ALIGNMENTS

RESULT 1
AX429465
LOCUS AX429465 2618 bp DNA linear PAT 21-UN-2002
DEFINITION Sequence 1 from Patent WO0234914.
ACCESSION AX429465
VERSION AX429465.1 GI:21540763
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Zhaio, Q., Beasley, E.M. and Wei, M.H.
Isolated human g-protein coupled receptors, nucleic acid molecules
encoding human gpcr proteins, and uses thereof
JOURNAL Patent: WO 0234914-A 1 02-MAY-2002;
PE CORP NY (US)
FEATURES
source
location/Qualifiers
1..2618
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 6.52e-137 Length: 2618
Score: 1763.00 Matches: 337
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x AX429465 (1-2618)

```

QY 1 MetGluSerLySerSerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIle 20
DB 447 ATGGAAAGCAATCTCTCATGGGTCATCAAGCTGGGTTCTGAGCATTCAGATTCACCATC 506
QY 21 ProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGluThrProCysTyrLys 40
DB 507 CCAGCTTGCGGTACAGAACTGACCAATCAACGAGAGGAGAGAGACTCTCTCTCAAG 566
QY 41 GlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsn 60
DB 567 CAGACCCTGAGCTTCACGGGGCTGAGCGTGCATGTTCCCTTCCGCTGACAGAAAC 626
QY 61 AlaValIleLeuThrPheLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIle 80
DB 627 GCGGTGTGCTCTGGCTCTCTGGGCTCCGATCGGAGAGAACCTGTCTTCATCTACATC 686
QY 81 LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyIleIleCysSerProLeu 100
DB 687 CTCAACCTGTCGCGGCGACCTTCTCTTCTTACGGGCGACATATATGTCGCGCTTA 746
QY 101 ArgLeuIleAsnIleArgHisProIleSerIleLeuSerProValMetThrPhePro 120
DB 747 CGCCCTCATCAATATCGGCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCC 806
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
DB 807 TACTTTATAGGCTTAAGCATGCTGAGCGCATCAGACGAGGCTGCTGCTCATCTC 866
QY 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
DB 867 TGGCCCATCTGGTACACATGCGCCGCCACAGATACCTGTCATCGGTGATGTGTC 926
QY 161 LeuThrAlaLeuSerLeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPhe 180
DB 927 CTCTGGGCGCTTCCTGCTGCGAGATCTGAGAGTGTGCTGTGATCTTCTGCT 986
QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
DB 987 AGTGTGTCGATCTCTCTTGGTGTGAACCTCAGATTCATTCATTCAGTGGTGTG 1046
QY 201 PheLeuCysValIleLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
DB 1047 TTTTATATGTCGCTCTCTGTCGAGTCAAGCTGCTGCTGTCAGGATTCCTGTGGA 1106
QY 221 SerArgLyMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValIleValPhe 240
DB 1107 TCCCGAAGATCCCGGTGACAGGCTGTACGTACCATCTCTCCACAGGCTGTGCTTC 1166
QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
DB 1167 CTCTCTGTGGGCTGCTCTTGGCATTCAGTGGGCTCTGTTCCAGATCCACTGGAT 1226
QY 261 TrpLysValIleLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
DB 1227 TGGAAAGTCTTATTTGTCATGTGATCTAGTTTCCATTTTCTGTGCGGCTCTTAACAG 1286
QY 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGln 300
DB 1287 AGTGCACCAACCCATCTTATCTTCCTGTGGGCTCTCTTAAGCAGGCTCAAAATAGAGCAG 1346
QY 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320
DB 1347 AACCTGAAGCTGGTTTCCAGAGGCTCTGACAGACACGCTTACAGTGAAGTGAAGTGA 1406
QY 321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
DB 1407 GGGTGGCTTCTCTACAGAAACCTGAGAGTGTGCGGAAAGCAGATTGGAGCAG 1457
  
```

RESULT 2
 AX429467
 LOCUS AX429467 8622 bp DNA linear PAT 21-JUN-2002
 DEFINITION Sequence 3 from Patent WO0234914.
 ACCESSION AX429467
 VERSION AX429467.1 GI:21540764
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Zhao, Q., Beasley, E. M. and Wei, M. H.
 Isolated human g-protein coupled receptors, nucleic acid molecules
 encoding human gpcr proteins, and uses thereof
 Patent: WO 0234914-A 3 02-MAY-2002;
 PE CORP NY (US)
 FEATURES
 source location/Qualifiers
 1..8622
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ALIGNMENT Scores:

Pred. No.: 2,366-133 Length: 8622
 Score: 1727.50 Matches: 332
 Percent Similarity: 96.81% Conservative: 2
 Best Local Similarity: 96.23% Mismatches: 2
 Query Match: 97.99% Indels: 9
 DB: 6 Gaps: 1

US-09-867-570-2 (1-337) x AX429467 (1-8622)

```

QY 2 GluSerLySerSerTrp-----ValIleArgLeuGly 12
DB 7459 CAGAGATCAACAGCGTGGATCAATCTGTTCTGTTCCAGGTCATCAGACTGGGG 7518
QY 13 PheLeuSerMetAspSerThrIleProValIleGlyThrGluLeuThrProIleAsnGly 32
DB 7519 TTTCTGAGCATGATTCACCATCCAGCTTGGGTGACAGACTGACACCAATCAACGGA 7578
QY 33 ArgGluGluThrProCysTyrLysGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
DB 7579 GGTGAGAGATCTCTTCTCAAGACAGACCTGAGCTTCACGGGCTGACGTGCTT 7638
QY 53 SerLeuValAlaLeuThrGlyAsnAlaValIleLeuThrLeuGlyCysArgMetArg 72
DB 7639 TCCCTGTGCGCCTGACAGAAACGGGTGTGCTCTGGCTCTGGGCTGCGCATGCGC 7698
QY 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSer 92
DB 7699 AGGAACGTCGTCTCCATCTACATCTCAACCTGTCGCGGCGGACCTTCTCTTACGC 7758
QY 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
DB 7759 GGCACATTATATGTCGCGGTACGCTTCATCAATATCCGACATCCATCTCCAAATC 7818
QY 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
DB 7819 CTCAGTCTGTATGATGACTTCTTCCATCTTATAGGCTTAAGCATCTAGAGCCATCAGC 7878
QY 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyr 152
DB 7879 ACCGAGCGTGCCTGTCATCTCTGTCGACCATCTGTACACTGCGCGCGCCAGATAC 7938
QY 153 LeuSerSerValMetCysValLeuLeuThrProIleAsnSerLeuLeuArgSerIleGlu 172
DB 7939 CTGTCACTGGATATGTCGCTCTGAGGCTCTGCTCTGCTCTGAGATCTCTGAG 7998
QY 173 TrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
  
```


Db 7999 TGGATGTTCTGACCTTCTGTTAGTGTGATTTCTGTTGTGTAAGCTGAT 8058
 QY 193 PheIleThrIleAlaTrpLeuValPheLeuCyseValleuCyseGlySerSerLeuVal 212
 Db 8059 TTCATTACATCGCGTGGCTGGTGTGTTTATGTTGTTCTGTGTGGGTCAAGCTGTCTC 8118
 QY 213 LeuLeuValAlaArgIleLeuCyseGlySerArgIleMetProLeuThrArgLeuTyValThr 232
 Db 8119 CTGCTGTGACGAGATCTCTGTGTGATCCCGAAGATGCCGTGACAGAGCTGTACGTACC 8178
 QY 233 IleLeuLeuThrValLeuValPheLeuLeuCyseGlyLeuProPheGlyIleGIntPrAla 252
 Db 8179 ATCTCTCTCAAGATCGCTGCTCTCTCTCTCTGTGGCTGTGCTTGGCATTCAGTGGCC 8238
 QY 253 LeuPheSerArgIleHisLeuAspTrpValLeuPheCyseHisValHisLeuValSer 272
 Db 8239 CTGTTTCCAGATCCACCTGATGTAAGTCTTATTTGTGATGTCATGATGTTTCC 8298
 QY 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrrPhePheValGlySer 292
 Db 8299 ATTTTCTCTGCTCCCTCTTAACAGACATGCCAACCCCACTTCACTTCTCTGTGGCTCC 8358
 QY 293 PheArgIleArgIleAsnArgIleAsnLeuLeuValLeuGlnArgAlaLeuGlnAsp 312
 Db 8359 TTTAGCAGCGCTCAAAATAGCAGAACCTGAGCTGTCTCCAGAGGCTCTGCAGGAC 8418
 QY 313 ThrProGluValAspGluGlyGlyIleTrpLeuProGlnIleuThrLeuGluLeuSerGly 332
 Db 8419 ACCCTGAGGTGATGAAAGGTGAGGGTGTCTCTCAGAAACCTCGAGCTGTCCGGA 8478
 QY 333 SerArgLeuGluGln 337
 Db 8479 AGCAGATTGGAGCAG 8493

RESULT 3
 LOCUS AX647081 1369 bp DNA linear PAT 04-MAR-2003
 DEFINITION Sequence 1273 from Patent EP1270724.
 ACCESSION AX647081
 VERSION AX647081.1 GI:28800068
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
 TITLE Guanine triphosphate-binding protein coupled receptors
 JOURNAL Patent: EP 1270724-A 1273 02-JAN-2003;
 National Institute of Advanced Industrial Science and Technology
 (JP) ; Center for Advanced Science and Technology Incubation, Ltd.
 (JP)

FEATURES
 source location/Qualifiers
 1..1369
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 201..1169
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAB69666.1"
 /db_xref="GI:28800069"
 /translation="MDSTIPVLGTELPINGREBTPCYKQTLSTFTGLTIVSLVALTG
 NAVLMLACGRMRNAVSIYILNLVADPFLSGHICSPRLINIRHPISKILSPVM
 TFPYFGLSMLSISTERCISLMPYICRPRPYLSWVCVLLMALSLRLSEMF
 CDPLFGANSVWCSTSPRITAMVPLCVLCCSLVLYLRICGSRKMLPTLYVTI
 LTLVTVFLCGLPRGIOMAFSRHIDMKVLPCHVHVSIFLSANSSANPITTFYFG
 SFRQRQRNQLKVLQALQDTPVDEGGMLPQETLISGRLEQ"

ORIGIN

Alignment Scores:
 Pred. No.:
 Score:

7.09e-134 Length: 1369
 1722.50 Matches: 331

Percent Similarity: 96.81% Conservative: 3
 Best Local Similarity: 95.94% Mismatches: 2
 Query Match: 97.70% Indels: 9
 DB: 6 Gaps: 1
 US-09-867-570-2 (1-337) x AX647081 (1-1369)

QY 2 GluSerIysSerTrp-----ValIleArgLeuGly 12
 Db 132 CAGAGATCAAAACAGCTGTGATGATCATCTGTTCTGTTCCAGGGCTACAGCTGGGG 191
 QY 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
 Db 192 TTTCTGAGCATGAGATTCACATCCAGTGTGGATGACAACTGACCAATCAACGGA 251
 QY 33 ArgGluIleuThrProCyseTyrlGlnThrLeuSerPheThrGlyLeuThrCyseIleVal 52
 Db 252 CGTAGGAGACTCTTGTCTACAGAGACCCTGAGCTTCAAGGGGCTGACGTGATGTT 311
 QY 53 SerLeuValAlaLeuThrGlyAsnAlaValValLeuTrpLeuLeuGlyCyseArgMetArg 72
 Db 312 TCCCTGTCCGCGCTGACAGGAAACGCGTGTCTCTGCTCTGCGCTCCGCGATCCG 371
 QY 73 ArgAsnAlaValSerIleTyrrIleLeuAsnLeuValAlaAlaAspPheLeuSer 92
 Db 372 AGGAGCTGTCTTCATCTACATCTTCACCTGATGCGGCCGACCTTCTCTCTTAGC 431
 QY 93 GlyHisIleIleCyseSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
 Db 432 GGCACATTATATGTTGCGCGTTACGCTCATCAATTCGGCATCCCATCTCCAAATC 491
 QY 113 LeuSerProValMetThrPheProTyrrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
 Db 492 CTCAGTCTGTGATGACCTTCCCTTATATAGGCTTAAGCATGTCAGACGCGCATCAGC 551
 QY 133 ThrGluArgCysLeuSerIleLeuTrpProIleTrpTyrrHisCysArgAspProArgTyrr 152
 Db 552 ACCGAGGCTGTCTGTCCATCTGTGGCCATCTGTGACCACTGACCGCCGCCCAAGATAC 611
 QY 153 LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuAspSerIleLeuGlu 172
 Db 612 CTGTCACTCATCATGATGTGTCTGTCTGTGGCCCTGTCCCTGTGCGGAGATATCTGGAG 671
 QY 173 TyrMetPheCysAspPheLeuPheSerGlyAlaAspSerValTyrrCysGluThrSerAsp 192
 Db 672 TGGATGTTCTGTGACTTCTGTTAGTGTGTCATATCTGTTGGTGTGAACCTGACAT 731
 QY 193 PheIleThrIleAlaTrpLeuValPheLeuCyseValValLeuCyseGlySerSerLeuVal 212
 Db 732 TTGATTACATCCGCGTGGCTGTTTTTATATGTTGTCTGTGTGGGTCCAGCTGTGTC 791
 QY 213 LeuLeuValAlaArgIleLeuCyseGlySerArgIleMetProLeuThrArgLeuTyValThr 232
 Db 792 CTGCTGTGACGATCTCTGTGTGATCCCGAAGATGCCGTGACAGGCTGTACGTACC 851
 QY 233 IleLeuLeuThrValLeuValPheLeuLeuCyseGlyLeuProPheGlyIleGIntPrAla 252
 Db 852 ATCTCTCTCAAGATGCTGCTCTCTCTCTGTGGCTGTGCTTGGCATTCAGTGGGCC 911
 QY 253 LeuPheSerArgIleHisLeuAspTrpValLeuPheCyseHisValHisLeuValSer 272
 Db 912 CTGTTTCCAGGATCCACCTGAGATGGAAGTCTTATTTGTGATGTCATCTAGTTTCC 971
 QY 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrrPhePheValGlySer 292
 Db 972 ATTTTCTCTGCTCCCTCTTAACAGACATGCCAACCCATCATTTACTTCTGTGGCTCC 1091
 QY 293 PheArgGlnArgIleAsnArgIleAsnLeuLeuValLeuGlnArgAlaLeuGlnAsp 312
 Db 1032 TTTAGGCAAGCTCAAAATAGCAGAACCTGAACTGCTTCTCCAGAGGGCTCTGCAGGAC 1091
 QY 313 ThrProGluValAspGluGlyGlyIleTrpLeuProGlnIleuThrLeuGluLeuSerGly 332

Db 1092 AGCGCTGAGTGGATGATGAGTGGAGGCTTCTCAGAGAACCTGAGAGTGTGGGA 1151
 Oy 333 SerArgLeuGln 337
 Db 1152 AGCAGATTGAGCAG 1166

RESULT 4
 AC108007/c
 LOCUS AC108007 91923 bp DNA 1linear PRI 03-SEP-2002
 DEFINITION Homo sapiens chromosome 11, clone CTD-3038L12, complete sequence.
 AC108007
 VERSION AC108007.5 GI:22655822
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Birren, B., Nussbaum, C., and Lander, E.
 TITLE Homo sapiens chromosome 11, clone CTD-3038L12
 JOURNAL Unpublished
 REFERENCE 2
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Glinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Karatas, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R.,
 Lander, T., Lehoczek, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McEwan, K., Meltrin, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunhang, P., Pierre, N., Pollard, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Lander, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meltrin, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 4
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Lander, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meltrin, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Sep 3, 2002 this sequence version replaced gi:22474954.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24596
 Center clone name: 3038_L_12

Only the first 91.9 kb of this clone are being submitted.
 The remainder overlaps accession number AC055860 (WICR project
 L9785).

FEATURES
 source

repeat_region
 25..70
 /rpt_family="L2"
 complement(251..899)
 repeat_region
 900..1001
 /rpt_family="L2"
 repeat_region
 1082..1655
 /rpt_family="L1MD2"
 repeat_region
 1671..1850
 /rpt_family="L1MD3"
 complement(1858..1917)
 repeat_region
 1918..2273
 /rpt_family="L2"
 repeat_region
 1918..2273
 /rpt_family="THE1B"
 complement(2274..2327)
 repeat_region
 4505..4713
 /rpt_family="L2"
 repeat_region
 5536..5795
 /rpt_family="L2"
 repeat_region
 5796..5837
 /rpt_family="AluSc"
 repeat_region
 5838..5875
 /rpt_family="TAAAn"
 repeat_region
 complement(7186..7526)
 /rpt_family="L1MB4"
 repeat_region
 complement(10716..10825)
 /rpt_family="MIR"
 repeat_region
 11369..11466

```

repeat_region /rpt_family="L2" complement(11479..11644)
repeat_region /rpt_family="MIR3" 12072..12409
repeat_region /rpt_family="MER2" 12692..12875
repeat_region /rpt_family="LTR33" complement(12876..13294)
repeat_region /rpt_family="MST1" 13295..13552
repeat_region /rpt_family="LTR33" complement(14619..14992)
repeat_region /rpt_family="L1PB1" complement(15002..17751)
repeat_region /rpt_family="L1PB1" 17752..19590
repeat_region /rpt_family="L1PA6" 19591..20184
repeat_region /rpt_family="L1PA4" 20185..24476
repeat_region /rpt_family="L1PA6" 24611..24637
repeat_region /rpt_family="(T)n" complement(24639..26213)
repeat_region /rpt_family="SVA" 26219..26361
repeat_region /rpt_family="L1MB5" 26372..26560
repeat_region /rpt_family="FLAM_C" complement(26636..26719)
repeat_region /rpt_family="MIR" complement(26863..27139)
repeat_region /rpt_family="MER41C" complement(27140..27449)
repeat_region /rpt_family="AluSq" complement(27450..27494)
repeat_region /rpt_family="MER41C" complement(27495..27796)
repeat_region /rpt_family="AluSg" complement(27797..28030)
repeat_region /rpt_family="MER41C" complement(28194..28287)
repeat_region /rpt_family="MIR" complement(28002..29304)
repeat_region /rpt_family="AluDo" complement(29337..29755)
repeat_region /rpt_family="MST1" 30412..30452
repeat_region /rpt_family="(TG)n" complement(31075..31197)
repeat_region /rpt_family="MIR" 31198..31246
repeat_region /rpt_family="(CA)n" complement(31247..31341)
repeat_region /rpt_family="MIR" 31541..31821
repeat_region /rpt_family="AluSx" 31822..31850
repeat_region /rpt_family="(GA)n" 32030..32230
repeat_region /rpt_family="MERSA" complement(32650..32747)

```

Alignment Scores:

```

Pred. No.: 9.88e-132 Length: 91923
Score: 1722.50 Matches: 331
Percent Similarity: 96.81% Conservative: 3
Best Local Similarity: 95.94% Mismatches: 2
Query Match: 97.70% Indels: 9
DB: 9 Gaps: 1

```

US-09-867-570-2 (1-337) x AC108007 (1-91923)

```

Qy 2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12
Db 9546 CAGAGATCAACAGCTGGTATGATCATCTGCTTTCTGTTCCAGGGTCATCAGCTGGG 9487
Qy 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
Db 9486 TTTCTGAGCATGATTCACATCCAGCTTGGGTACAGAACTGACCATCAAGCGGA 9427
Qy 33 ArgGluGluThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 9426 CGTAGAGAGACTCTGCTCAAGAGAGACCTGAGCTTCAAGGGGCTGAGGTGATGTT 9367
Qy 53 SerLeuValIleLeuThrGlyAsnAlaValIleLeuThrLeuLeuGlyCysArgMetArg 72
Db 9366 TCCCTGTCGCTGAGAGAGAAAGCGGTGTGCTGAGCTCTGAGGCTCCGATGCGC 9307
Qy 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuSer 92
Db 9306 AGGAGCGCTGTCTCCATCTACATCTCCACCTGCGCGGCGCATCTCTCTTCC 9247
Qy 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
Db 9246 GGCACATATATATGTTGCGCGTTACGCTCATCAATCCGATCCATCTCCAAATC 9187
Qy 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
Db 9186 CTCAGTCCGTGATGATGACCTTCCCTACTTATAGGCTTAAGCATGTCGAGCGCATCAGC 9127
Qy 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrPheHisCysArgGlyProArgTyr 152
Db 9126 ACCGAGCGCTGCTGCTTCATCTCTGTGCGCATCTGATACATGCGCCGCCCAAGTAC 9067
Qy 153 LeuSerSerValMetCysValLeuLeuThrIleLeuSerLeuLeuArgSerIleLeuGlu 172
Db 9066 CTGTCATCGGTATGATGTGTGCTGCTGAGCGCTGCTGCTGCGGAGATATCTGGAG 9007
Qy 173 TyrMetPheCysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAsp 192
Db 9006 TGAATGTTCTGTACTCTCTGTTAGTGATGATTCATTCGTTGTGTGAAGCTCATGAT 8947
Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuVal 212
Db 8946 TTGATTCATATCCGCTGCTGCTGCTTTTATATGTGTCTCTGTGGGTCAACCTGTC 8887
Qy 213 LeuLeuValArgIleLeuCysGlySerArgLysMetProLeuThrArgLeuTyrValThr 232
Db 8886 CTGCTGCTCAGATTCCTGTGTGATCCGGAAGATGCGCTGACAGGCTGTACGTACC 8827
Qy 233 IleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAla 252
Db 8826 ATCTCTCTCAGATGCTGCTCTCTCTCTGAGGCTGCTTGTGATTCAGTGGGCC 8767
Qy 253 LeuPheSerArgIleHisLeuAspTrpLysValLeuPheCysHisValHisLeuValSer 272
Db 8766 CTGTTTTCAGAGATCACTGATGATGAAAGTCTTATTTGTGATGTCATCTAGTTTCC 8707
Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySer 292
Db 8706 ATTTCTGTCGCTCTTAACAGACAGTCCAACTCATATTTACTTCTCGTGGGCTCC 8647
Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 8646 TTTAGCAGCGTCAAAATAGGCAGAACTGAACTGTTCTCCAGAGGGCTGCGCAGAC 8587
Qy 313 ThrProGluValAspGluGlyGlyTyrPheLeuProGluGlnThrLeuGluLeuSerGly 332
Db 8586 ACCCTCGAGGTGATTAAGGTGAGGGTGTCTCTCAGAAACCTTGAGCTGTGCGGA 8527
Qy 333 SerArgLeuGluGln 337
Db 8526 AGCAGATTGAGCAG 8512

```

RESULT 5

AC090099	172939 bp	DNA	linear	PR1 01-OCT-2002
LOCUS				
DEFINITION	Homo sapiens chromosome 11, clone RP11-113D6, complete sequence.			
ACCESSION	AC090099	GI:23396332		
VERSION	AC090099.14			
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 172939)			
AUTHORS	Birren, B., Nusbaum, C., and Lander, E.			
TITLE	Homo sapiens chromosome 11, clone RP11-113D6			
JOURNAL	Unpublished			
REFERENCES	2 (bases 1 to 172939)			
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Heath, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatae, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Souner, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Trivis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 172939)			
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatae, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	4 (bases 1 to 172939)			
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatae, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	On Oct 1, 2002 this sequence version replaced gi:22726159.			
AUTHORS	All repeats were identified using RepeatMasker: Smt, A.F.A. & Green, P. (1996-1997)			
TITLE	http://ftp.genome.washington.edu/RM/RepeatMasker.html			
JOURNAL	Direct Submission			
COMMENT	Center: Whitehead Institute/ MIT Center for Genome Research Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu Center project name: 113_D_6 Center clone name: 113_D_6			
FEATURES	Location/Qualifiers			
source	1..172939			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	/chromosome="11"			
	/map="11"			
	/clone="RP11-113D6"			
	/clone_1lb="RP11-11 Human Male BAC"			
repeat_region	complement(1..77)			
	/rpt_family="Alus"			
repeat_region	complement(154..680)			
	/rpt_family="Char1el"			
repeat_region	complement(678..810)			
	/rpt_family="Char1el2"			
repeat_region	complement(879..1057)			
	/rpt_family="L1M4"			
repeat_region	complement(1395..1513)			
	/rpt_family="FLM_A"			
repeat_region	complement(1876..1876)			
	/note="<30 qual SNGL region"			
repeat_region	complement(1902..1910)			
	/note="<30 qual SNGL region"			
repeat_region	complement(2672..2788)			
	/rpt_family="S5"			
repeat_region	complement(4009..4170)			
	/rpt_family="MIR"			
repeat_region	complement(4210..4355)			
	/rpt_family="MER5B"			
repeat_region	complement(4356..4647)			
	/rpt_family="AlusX"			
repeat_region	complement(6554..6658)			
	/rpt_family="LRR41"			
repeat_region	complement(6916..6985)			
	/rpt_family="MLT1C"			
repeat_region	complement(7175..7359)			
	/rpt_family="LRR41"			
repeat_region	complement(8308..8611)			
	/rpt_family="AluJb"			
repeat_region	complement(8622..9083)			
	/rpt_family="MLT2B1"			
repeat_region	complement(9265..10668)			
	/rpt_family="L1PB3"			
repeat_region	complement(10699..10809)			
	/rpt_family="MER34"			
repeat_region	complement(10782..11264)			
	/rpt_family="LRR48B"			
repeat_region	complement(11344..11683)			
	/rpt_family="L2"			
repeat_region	complement(11839..12026)			

```

repeat_region      /rpt_family="L1MC4a"
12188..12660
/rpt_family="L1MA2"
repeat_region      /rpt_family="L1MA2"
12879..13222
/rpt_family="L1MC4a"
repeat_region      /rpt_family="L1MC4a"
13223..13522
/rpt_family="AlusX"
repeat_region      /rpt_family="L1MC4a"
13523..13817
/rpt_family="L1MC4a"
repeat_region      /rpt_family="AluJo"
14819..15108
/rpt_family="AluJo"
repeat_region      /rpt_family="AluJo"
15114..15175
/rpt_family="(GAA)n"
repeat_region      /rpt_family="MERSA"
16150..16336
/rpt_family="MERSA"
/rpt_family="MERSA"
complement(16587..16598)
/note="<30 qual SNGL region"
complement(16607..16612)
/note="<30 qual SNGL region"
complement(16746..17360)
/rpt_family="L1ME3"
repeat_region      /rpt_family="AluJo"
17361..17664
complement(17665..17774)
/rpt_family="L1ME3"
repeat_region      /rpt_family="L1ME3"
complement(17827..17929)
/rpt_family="MIR"
repeat_region      /rpt_family="MIR"
18993..19298
/rpt_family="AluJo"
repeat_region      /rpt_family="AluJo"
19759..19821
/rpt_family="L2"
repeat_region      /rpt_family="L2"
19933..19941
/note="<30 qual SNGL region"
repeat_region      /rpt_family="AluJo"
20320..20609
complement(20684..20994)
/rpt_family="AlusX"
repeat_region      /rpt_family="AlusX"
20713..20721
/note="<30 qual SNGL region"
repeat_region      /rpt_family="AluJo"
20879..20883
complement(21116..21398)
/rpt_family="AlusX"
repeat_region      /rpt_family="AlusX"
21141..21147
/note="<30 qual SNGL region"
repeat_region      /rpt_family="AlusX"
21144
/note="<30 qual SNGL region"
repeat_region      /rpt_family="AlusX"
21146
/note="probably A; possibly G"
repeat_region      /rpt_family="AlusX"
21153..21158
/note="<30 qual SNGL region"
repeat_region      /rpt_family="AlusX"
21154
/note="probably T; possibly C"
repeat_region      /rpt_family="AlusX"
21160..21167
/note="<30 qual SNGL region"

```

Alignment Scores:

```

Pred. No.:      2.07e-131      Length:      172939
Score:          1722.50      Matches:      331
Percent Similarity: 96.81%      Conservative: 3
Best Local Similarity: 95.94%      Mismatches: 2
Query Match:      97.70%      Indels:      9
DB:              Gaps:      1

```

US-09-867-570-2 (1-337) x AC090099 (1-172939)

```

QY      2 GluseryserTP-----ValleatgLeugly 12
DB      114041 CACAGATCAAAAGCTGTATCATCTGTTCTGTTCCAGGGTCATCAGACTGGGG 113982
QY      13 PheluserTPserTPThrlleProvalleuJlThrlleuTPProilLeangly 32
DB      113981 TTTCTAGCATGATTCACCATCCAGCTCTGGGTACAGAACTGACCAATCAACGGA 113922

```

```

QY      33 ArgGluGluThrProCyTrlysgLInThrLeuserPheThrlleuThrCyVelleVal 52
DB      113921 CGTAGAGAGACTCCTGTCTACAGAGACAGACCTGAGCTTCCAGGGGCTGAGAGCTGCTT 113862
QY      53 SerLeuValAlaLeuThrGlyAsnAlaValLeuThrLeuLeuGlyCyAspMetArg 72
DB      113861 TCCCTTGCGCTGACAGAGAAACGGGTGTGTCTCTGCTCTCTGAGCTGCGCATGCGC 113802
QY      73 ArgAsnAlaValSerIleThrlleAsnLeuValAlaAspPheLeuPheLeuSer 92
DB      113801 AGGAAGCTGTCTCCATCATCAACCTCAACCTGTGCGCGCAGCTTCTCTCTTACG 113742
QY      93 GlyHleIleLeuCySerProLeuArgLeuIleAsnIleArgHisProIleSerlyIle 112
DB      113741 GGCCACATATATATGTTGCGCGTTACCCCTCATCAATATCCGCACTCCCAATATC 113682
QY      113 LeuserProValMetThrPheProTyrrPheIleGlyLeuserMetLeuserAlaIleSer 132
DB      113681 CTGAGTCTGTGTATGACCTTCTCTCTTATAGCCCTAAGCTAAGCTGAGCGGCATCAGC 113622
QY      133 ThrGluArgCyLeuserIleLeuThrProIleThrlleCyAspArgProArgTyr 152
DB      113621 ACCGACGCTGCTGTCTCATCACTGTGCGCATCTGTACACACGCGCGCCAGATAC 113562
QY      153 LeuserSerValMetCyValLeuLeuThrPalaLeuserLeuLeuArgSerIleLeuGlu 172
DB      113561 CTGTCAATCATGATCATGTGTGTCCTGCTGCGGCTCTCTCTCTGCGGATATCTCGAG 113502
QY      173 TrpMetPheCyAspPheLeuPheSerGlyAlaAspSerValTrpCyGluThrSerAsp 192
DB      113501 TGGATGTCTGTACTCTCTGTTATGAGTGCTAATCTGTGTGTGTGAAGCTCATGAT 113442
QY      193 PheIleThrlleatThrlleuValPheLeuCyValValLeuCyGlySerSerIleVal 212
DB      113441 TTCAATTCATCAATCCGCGCTGCTGTTTATATGTGTCTCTGTGCGTCAACCTGCTC 113382
QY      213 LeuLeuValArgIleLeuCyGlySerArgLysMetProLeuThrArgLeuIleValThr 232
DB      113381 CTGCTGTGTCAGATTCCTGTGTGATCCGGAATGCGCGTGCACAGGCTGTAGTACC 113322
QY      233 IleLeuLeuThrValLeuValPheLeuLeuCyGlyLeuProPheGlyIleGlnTrpAla 252
DB      113321 ATCCCTCTCAAGATGTGCTCTCTCTCTGTCGCTGCTGTCATTCAGTGGGCC 113262
QY      253 LeuPheSerArgIleHisLeuAspTrpLysValLeuPheCyHisValHisLeuValSer 272
DB      113261 CTGTTTCCAGGATTCACCTGATGAGATGAGAACTTATTTGTATGATGATCTAGTTCC 113202
QY      273 IlePheLeuserAlaLeuAsnSerSerAlaAsnProIleIleThrllePheValGlySer 292
DB      113201 ATTTCTGTCCTGCTCTTACACAGATGCAACCCCATCATTTACTTCTGAGGCTCC 113142
QY      293 PheArgGlnArgGlnAsnAlaGlnAsnLeuLeuValLeuGlnArgAlaLeuGlnAsp 312
DB      113141 TTATAGCAGCGTCAAAATAGACGAACCTGAACCTGTTCTCCAGAGGGCTGACAGAC 113082
QY      313 ThrProGluValAspGlyGlyGlyThrlleProGlnIleThrlleuGluLeuSerGly 332
DB      113081 ACCGCTGAGGTGATGAAGGTGAGGGTGTCTCTCAGAAACCTGAGAGCTGTGGGA 113022
QY      333 SerArgLeuGluGln 337
DB      113021 AGCAGATTGAGCAG 113007

```

RESULT 6

```

AX299158      AX299158      1400 bp      DNA      linear      PAT 26-NOV-2001
LOCUS          Sequence 30 from Patent WO0183555.
ACCESSION      AX299158
VERSION        AX299158.1 GI:17129065
KEYWORDS       Homo sapiens (human)
SOURCE

```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Anderson, D.J., Dong, X., Zylka, M., Han, S.K. and Simon, M.
Pain signaling molecules
Patent: WO 0183555-A 30 08-NOV-2001;
JOURNAL CALIFORNIA INSTITUTE OF TECHNOLOGY (US)

FEATURES

Location/Qualifiers

CDS

1..1400
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
332..1330
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD12713.1"
/db_xref="GI:17129066"
/translation="MDSTIPVLGTELTPINGREBPCKQTSFTGLTCTISVALTG
NAVAVMLGCRMRNRNAYSIYILNVLADLPFSGHICSPRLINIRHPSKILSEPM
TEPYIGLSMSAISSTERCLSIIMPYHCRPRYSYVCVLNLSLSLEWFM
CDPLFSGADSVWCSTEDPITIAMVPLCYVLGSSSLVILVRLICGRKMLRLVYTI
LLTVVPLICGPRFGIOWALFRIHLDKRYLPCRYHVLVIFISALNSANPILITFVVG
SFRQRKRNKMLVLPALQDTPVEDSGGMLPQETLRLSGSRLEQ"

ORIGIN

Alignment Scores:

Pred. No.: 9.7e-134 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x AX299158 (1-1400)

QY 6 SerTPValIIleArgLeuGIyPheLeuSerMetAspSerThrIleProValLeuGIyThr 25
DB 302 TCACAGGCTACACAGCTGGGTTTCTGACATGATTCACACATCCCAAGCTTGGGTAC 361
QY 26 GluLeuThrProIIleAsnGIyArgGIuGIuThrProCyGlyTyrIleGlnThrLeuSerPhe 45
DB 362 GAACGACACCAATCAACGACGAGTGAAGACCTCTGCTACCAAGACACCTGAGCTTC 421
QY 46 ThrGIyLeuThrCySIIleValSerLeuValAlaLeuThrGIyAsnAlaValIleLeuTP 65
DB 422 ACGGGGCTGACGCTGATCTTCCCTGTCGCCCTACAGAAACCGGTTGGCTCTG 481
QY 66 LeuLeuGIyCyArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
DB 482 CTCCTGGGCTGCGCATGCGACGAAAGCTGTCTCATCTACACCTGCTGCGG 541
QY 86 AlaAspPheLeuPheLeuSerGIyHisIleIleCySerProLeuArgLeuIleAsnIle 105
DB 542 GCCGACTTCCTCTTCTTACCGGCCCATATATGTTCCGCTTACGCTCATCATATTC 601
QY 106 ArgHisProIIleSerIyIleLeuSerProValMetThrPheProTyrPheIleGIyLeu 125
DB 602 CGCCATCCCATCTCAAAATCTCTCACTCTGTGATACACCTTCCCTTATAGGCTTA 661
QY 126 SerMetLeuSerAlaIleSerThrGIyArgCySLeuSerIleLeuThrProIleTyrIy 145
DB 662 AGCATGTGAGCGCATGACACCGACGCGCTGCTCATCTGCGCCCATCTGCTAC 721
QY 146 HisCyArgArgProArgTyrIleuSerSerValMetCyValLeuLeuThrAlaLeuSer 165
DB 722 CACTCCCGCGCCCAAGATACCTGTCACTGATATGTCTGTCTGCGGCCCTGTCC 781
QY 166 LeuLeuArgSerIleLeuGIuThrMetPheCyAspPheLeuPheSerGIyAlaAspSer 185
DB 782 CTGCTGCGGAGATCTCTGAGATGATCTTCTGTGATCTTCTGTGATGCTGATTTCT 841
QY 186 ValTrpCySgluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCySValVal 205

DB 842 GTTGTGTGAAGACGACATTTCAATTCAGCGGTGCTGTTTATATGTGTGTT 901
QY 206 LeuCyGIySerSerLeuValLeuLeuValArgIleLeuCyGIySerArgIySerPro 225
DB 902 CTCGTGGGTTCAGCGCTGTCTGCTGTGACAGATTCCTGTGATCCCGGAAGATGCG 961
QY 226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCySLeu 245
DB 962 CTGACAGGCTGTACGTGACATCTCTCCACAGAGCTGTGTTCTCTCTGTGTGCTG 1021
QY 246 PropheGIyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIyValIleuPhe 265
DB 1022 CCTTTGGCATTCAGTGGGCGCTGTTCACAGATCCACTGATTTGGAAGATCTATTT 1081
QY 266 CyHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerAlaAspProIle 285
DB 1082 TGTCAATGTCATCTAGTTTCCATTTTCTGTCCGCTTTAACAGCAGTCCAACTCCATTC 1141
QY 286 IleTyrPhePheValGIySerPheArgGlnArgGlnAsnArgGlnAsnLeuIyLeuVal 305
DB 1142 ATTACTTCTTGTGTGGCTCTTTAGGACGCTCAAAATAGGCAGAACCTGAAGCTGTT 1201
QY 306 LeuGlnArgAlaLeuGlnAspThrProGIuValAspGIuGIyGIyTyrLeuProGln 325
DB 1202 CTCACAGGCTCTGTCAGAGACAGCGCTGAGGTGATGAAGGTGAGGCTGCTCTCAG 1261
QY 326 GluThrLeuGIuLeuSerGIySerArgLeuGIuGln 337
DB 1262 GAACCTGTGACCTGTGCGGAAAGATTTGAGCAG 1297
RESULT 7
AX549388 1400 bp DNA linear PAT 26-NOV-2002
LOCUS AX549388
DEFINITION Sequence 673 from Patent WO02061087.
ACCESSION AX549388
VERSION AX549388.1 GI:25814004
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Burner, G.C., Roub, C.L. and Brown, J.P.
Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
Patent: WO 02061087-A 673 08-AUG-2002;
JOURNAL Lifespan Biosciences, Inc. (US)
FEATURES
source
1..1400
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Alignment Scores:
Pred. No.: 9.7e-134 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 6 Gaps: 0
US-09-867-570-2 (1-337) x AX549388 (1-1400)
QY 6 SerTPValIIleArgLeuGIyPheLeuSerMetAspSerThrIleProValLeuGIyThr 25
DB 302 TCACAGGCTACACAGCTGGGTTTCTGACATGATTCACACATCCCAAGCTTGGGTAC 361
QY 26 GluLeuThrProIIleAsnGIyArgGIuGIuThrProCyGlyTyrIleGlnThrLeuSerPhe 45
DB 362 GAACGACACCAATCAACGACGAGTGAAGACCTCTGCTACCAAGACCTGAGCTTC 421

QY 46 ThGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyValAlaValLeuTyr 65
 DB 422 ACCGGGGCTGACGGACGCGTTTCTTGTGCGCGTGAACAGAAACGGGTTGCTCTGG 481
 QY 66 LeuLeuGlyCYsArGMeTArGArGAnaAlaValSerIleTyrIleLeuAuaLeuValAla 85
 DB 482 CTCTGGGGCTGGCGACATGCGAGAGAGCGTGTCTCAATCAATCTCAACCTGGTGGCG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
 DB 542 GCCGACTTCTCTCTCTTCTTACGGCGCAATTATATGTTGGCGTTAGCCCTCAATATC 601
 QY 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
 DB 602 CGCCATCCCAATCCCAAAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAAT 661
 QY 126 SerMetLeuSerAlaIleSerThrGlyArgCysLeuSerIleLeuTyrProIleTyr 145
 DB 662 AGCATGCTGAGCGCCATCAACAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
 QY 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTyrAlaLeuSer 165
 DB 722 CACTGCG 781
 QY 166 LeuLeuArgSerIleLeuGlyTyrMetPheCysPhePheLeuPheSerGlyAlaAspSer 185
 DB 782 CTCTCGCGGAGTATCTCGGAGTATCTCGGAGTATCTCGGAGTATCTCGGAGTATCT 841
 QY 186 ValTyrCysGlyLeuThrSerAspPheIleThrIleAlaTyrLeuValPheLeuCysValVal 205
 DB 842 GTTGGTGGTGAACGTCAGATTCATTACATGCGGCGGCTTTTATATGAGTGT 901
 QY 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
 DB 902 CTCTGAGGCTGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
 QY 226 LeuThrArgLeuTyrValIleThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
 DB 962 CTACACGAGCGCTGACGACATCTCTCAAGTGTGCTCTCTCTCTCTCTCTCTCTCT 1021
 QY 246 ProPheGlyIleGlnTyrAlaLeuPheSerArgIleHisLeuAspTyrPheValLeuPhe 265
 DB 1022 CCTTTGGCATGAGTGGGCGCTGTTTCCAGATCCACCTGGATGGAAAGCTTATTT 1081
 QY 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 DB 1082 TGTCAATGATCTAGTTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
 QY 286 IleTyrPhePheValGlySerPheArgIleArgIleAsnArgIleAsnLeuValLeuVal 305
 DB 1142 ATTACT 1201
 QY 306 LeuGlyArgAlaLeuGlnAspThrProGlyValAlaAspGlyGlyTyrTyrLeuProGln 325
 DB 1202 CTCAGAGGGCTCTGAG 1261
 QY 326 GluThrLeuGlyLeuSerGlySerArgLeuGlyGln 337
 DB 1262 GAAACCTGAGGCTGCTGCGGAGAGATTGGAGAG 1297
 RESULT 8
 AY042215 1400 bp DNA linear PRI 14-SEP-2001
 LOCUS Homo sapiens G protein-coupled receptor (MRGX3) gene, complete cds.
 ACCESSION AY042215
 VERSION AY042215.1 GI:15546065
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1400)

AUTHORS Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
 TITLE A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons
 JOURNAL Cell 106 (5), 619-632 (2001)
 MEDLINE 21435808
 PUBMED 11551509
 REFERENCE 2 (bases 1 to 1400)
 AUTHORS Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
 TITLE Direct Substitution
 JOURNAL Submitted (19-JUN-2001) Division of Biology, 216-76, California Institute of Technology, 1201 E. California Blvd., Pasadena, CA 91125, USA
 FEATURES
 source Location/Qualifiers
 1..1400
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <332..>1300
 /gene="MRGX3"
 <332..>1300
 /gene="MRGX3"
 /product="G protein-coupled receptor"
 332..1300
 /gene="MRGX3"
 /note="Mas-related gene X3"
 /codon_start=1
 /product="G protein-coupled receptor"
 /protein_id="AK91806.1"
 /db_xref="GI:15546066"
 /translation="MDSTIPVLTETLPINGRETPCYKOTLSFTGTCIVSLALNG
 NAYVIMLGLGRNRNAYSTIYNLVADPLFSGHICSTRLINIRPISLSPM
 TPEYFGLSLAISTERCSTLPIWHORRRYSSVCVLMALSLRSLDEMF
 CDPLFSGADSWCETSDFTIAMLVFLCVLFCSSSLVLRILCGSRKMLPLRYVI
 LRLVFLPLCGLPFGIOMALFSRIHDMKVLFCVHIVSIFLSALNSANPIYFVG
 SFRQNRQNRQLVLRALQDTPREVDEGGMLPEFTLELSRLQG"
 ORIGIN
 Alignment Scores:
 Pred. No.: 9, 7e-134 Length: 1400
 Score: 1721.00 Matches: 330
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 97.62% Indels: 0
 DB: Gaps: 0
 US-09-867-570-2 (1-337) x AY042215 (1-1400)
 QY 6 SerTyrValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 DB 302 TCCAGGGGTCCAGACTGGGGGTTTCTGAGCATGATTCACCAATCCAGTCTGGGTACA 361
 QY 26 GluLeuThrProIleAsnGlyArgGlyGlnTyrProCysTyrIleGlnThrLeuSerPhe 45
 DB 362 GAATGACACCAATCAACGAGACCTGAGAGACTCTTGTCTACACAGACGACCTGAGCTTC 421
 QY 46 ThGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyValAlaValLeuTyr 65
 DB 422 ACCGGGGCTGACGGACGCGTTTCTTGTGCGCGTGAACAGAAACGGGTTGCTCTGG 481
 QY 66 LeuLeuGlyCYsArGMeTArGArGAnaAlaValSerIleTyrIleLeuAuaLeuValAla 85
 DB 482 CTCTGGGGCTGGCGACATGCGAGAGAGCGTGTCTCAATCAATCTCAACCTGGTGGCG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
 DB 542 GCCGACTTCTCTCTCTTCTTACGGCGCAATTATATGTTGGCGTTAGCCCTCAATATC 601
 QY 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
 DB 602 CGCCATCCCAATCCCAAAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAAT 661
 QY 126 SerMetLeuSerAlaIleSerThrGlyArgCysLeuSerIleLeuTyrProIleTyr 145

Db	662	AGCAGTCGAGGCGCATCAGACCGAGGCGTCGTCGTCATCTGTGGCCATCTGATC	722
Qy	146	HisCybaTgaTgPrOiaTgTyrLeuSerSerValMetCyValLeuLeuTPrAlaLeuSer	165
Db	722	CACGCCCCCGCCCGCAGATACCTGTCATCGGTCATGTGTCCTCTCTGGGCCCTGTGCC	781
Qy	166	LeuLeuAgsSerIleLeuGluTTrpPheCyAspPheLeuPheSerGlyValAspSer	185
Db	782	CTGCTGCGGAAGTATCTTGAGTGGATGTTCTGTGACTTCTGTTTAGTGGTGAATCT	841
Qy	186	ValTPrCySgluThrSerAspPheIleThrIleAlaTPrLeuValPheLeuCyValVal	205
Db	842	GTTTGGTGTGAACGTCAGATTTCATTCAATCGGTGGCGTGTGTTTTTATGTGTGTT	901
Qy	206	LeuCySglSerSerLeuValLeuLeuValAlaGlyIleLeuCySglSerArgLeuMet	225
Db	902	CTCTGTGGGTCCAGCGCTGTGCTCTGTGTCAGAGATTCTCTGTGATCCCGGAAAGTGGCG	961
Qy	226	LeuTPrAgsLeuTyrValThrIleLeuLeuThrValIleValPheLeuLeuCySglLeu	245
Db	962	CTGACCAAGCTGTGATCGTACACCAATCTCTCACAGTCTGGGTCTTCTCTGTGGCGCTG	1022
Qy	246	ProPheGlyIleGlnTPrAlaLeuPheSerArgIleHisLeuAspTPrLySerValLeuPhe	265
Db	1022	CCCTTTGGCATTCAGGTGGGCCCTGTTTTCAGAGATCCAGCTGGAATGGAAAGTCTTAATT	1081
Qy	266	CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle	285
Db	1082	TGTCAATGTGCAATCTGATTTCCATTTTCCGTGCTCTTAACAGACGTGCCAACCCCATTC	1141
Qy	286	IleTTrPhePheValGlySerPheArgGlnAlaGlyAlaAsnArgGlnAsnLeuYseVal	305
Db	1142	ATTTACTTCTTGTTGGGCTCTTTTAGCGAGCTCAAAATAGGCAAACTGAAGCTGGTT	1201
Qy	306	LeuGlnArgAlaLeuGlnAspThrProGlnValAlaGlyGlyGlyGlyTTrPheProGln	325
Db	1202	CTCCAGAGAGGCTCTCAGGACACGCGCTGAGTGGATGAAGTGGAGGGGTCTTCTCAG	1261
Qy	326	GluThrLeuGluLeuSerGlySerArgLeuGluGln	337
Db	1262	GAACCCCTGAGCTGTCCGGAAGCAGATTGGAGCAG	1297
RESULT 9			
LOCUS	BC067292	1697 bp	mRNA
DEFINITION			linear
ACCESSION			BC067292
VERSION			1
KEYWORDS			MG.1
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
			1 (bases 1 to 1697)
			Strausberg,R.L., Feingold,E.A., Gronow,L.H., Derge,J.G.,
			Klausner,R.D., Collins,F.S., Wagner,L., Schmen,C.M., Schuler,G.D.,
			Altschul,S.F., Zeeberg,B., Buecwo,H., Scheffer,C.F., Bhat,N.K.,
			Hopkins,R.D., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
			Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
			Stepleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
			Scheer,I.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
			Carinini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
			Abrahamson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
			McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
			Worley,D.C., Hale,S., Garcia,A.M., Gay,L.J., Hilyk,S.W.,
			Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
			Fahney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
			Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
			Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
			Dickson,M.C., Rodriguez,A.C., Gilmwood,T., Schmutz,J., Myers,R.M.,
			Butterfield,Y.S., Kravynetski,M.I., Skalek,U., Smallus,D.E.,
			Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1697)
 Strauberg, R.
 Direct Submission
 Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Tekla Olson, Diana Palquist, Anna Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelica Schercher, Ursula Skalka, Duane Smallin, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES
 source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 141 Row: m Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 16876452
 Location/Qualifiers
 1..1697
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:75450 IMAGE:30383708"
 /tissue_type="Testis"
 /clone_id="NIH_MGC_180"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6.1"
 1..1697
 /gene="MRGX3"
 /db_xref="locusID:117195"
 /db_xref="MIM:607229"
 362..1330
 /gene="MRGX3"
 /codon_start=1
 /product="g protein-coupled receptor MRGX3"
 /protein_id="AAH67292.1"
 /db_xref="GI:45501077"
 /db_xref="locusID:117195"
 /db_xref="MIM:607229"
 /translation="MDSITVLTGTELPINGREBTPCYKQTLSFTGLCTIVSLVATLGG
 NAAVVLMLGCMRRNNAVSITVLNVADPFLSGHIIKSPRLINIRIPKSLSPVW
 TPFPIGIGSMLSAISTRCSLIWPPIWHCRPRVLSVMCVLMALSLASILEMMF
 CDPFGSGDSVWCETSDPFTIAMLVPLCVGSGSLVLLVILCGSRKMPRLVYTI
 LITATVILVLTGSLPGIOMALFSRTHDMKVLFCVHVLSVIFSLANSANPIIYFVGG
 SFRORNRNKLKVLQRLADPTPEVDEGGGLPQETLELSGRLEQ"

ORIGIN

Alignment Scores:
 Pred. No.: 6.77e-132 Length: 1697
 Score: 1700.00 Matches: 328
 Percent Similarity: 99.39% Conservative: 0
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 96.43% Indels: 0
 Gaps: 0 DB: 0

US-09-867-570-2 (1-337) x BC067292 (1-1697)


```

QY      8 ValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeu 27
Db      338 GTATCATGAGTGGAGGTTCTGAGACGAGATTCACCACTCCAGCTTGGGTACGAACTG 397
QY      28 ThrProIleAsnGlyArgGluGluThrProCysTyrIleGlnIleThrLeuSerPheThrGly 47
Db      398 ACACCAATCAACGAGACCTGAGAGACTCTTGTCTACAGACGACCTGAGACTTCACGGGG 457
QY      48 LeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValAlaValLeuThrLeu 67
Db      458 CTGACGCGCATCGTTCTCTTGTCTGCGCGCTGACAGAAACCGGTTGTGCTCTGCTCTG 517
QY      68 GlyCysArgMetArgAspAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAsp 87
Db      518 GGCTGCGCAGATGCGCAGAGAACGCTGTCTCATCTCAACCTCGTGGCGGCGGAC 577
QY      88 PheLeuPheLeuSerGlyHisIleCysSerProLeuArgLeuIleAsnIleArgHis 107
Db      578 TTCTCTTCTTCTTACGGCCCATTTATACGTTCCGCGTTACGCTCATCAATATCCGCAT 637
QY      108 ProIleSerIleValLeuSerProValMetThrPheProTyrPheIleGlyLeuSerMet 127
Db      638 CCCATCTCCAAATTCCTCAGTCCGTGATGACCTTCCCTACTTATAGCGCTAAGCATG 697
QY      128 LeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleTyrIleCys 147
Db      698 CTGAGGCGCCATCGACACCGAGCGCTGCTGTCCATCTGTGGCCCATCTGTATCCACTGC 757
QY      148 ArgArgProArgTyrLeuSerSerValMetCysValLeuLeuThrAlaLeuSerLeu 167
Db      758 CGCGCGCCCGAGATACCTGTATCGGTATGTGTCTGTGCGCGCTGTCTCTCTG 817
QY      168 ArgSerIleLeuGluThrPheCysAspPheLeuPheSerGlyAlaAspSerValTyr 187
Db      818 CGAGATATCTGAGATGATGTTCTGTGACTTCTGTGTTAGTGATGCTGATTCGTTGG 877
QY      188 CysGluThrSerPheIleThrIleAlaThrLeuValPheLeuCysValValLeuCys 207
Db      878 TGTGAAACGTCAATTTTCAATTCATTCAGTCGTGTGTATTTATGTGTGCTCTGT 937
QY      208 GlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetProLeuThr 227
Db      938 GGGTTCAGCTGTGCTCTGTGTCAGGATTCCTGTGATTCGCGAAAGATGCGCTACCC 997
QY      228 ArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPhe 247
Db      998 AGGCTGAGCTGACCACTCTCTCACAGTCTGTCTCTCTCTGTGCGCTGCTCTT 1057
QY      248 GlyIleGlnThrAlaLeuPheSerArgIleHisLeuAspTyrIleValLeuPheCysHis 267
Db      1058 GGCATTTCAGTGGGCGCTGTCTTCCAGGATCCACTGTGATGGAAAGCTTATTTTGTCT 1117
QY      268 ValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyr 287
Db      1118 GTTCATCTATATTCATTTCCTGTCGCTTTCATTCACGACGTCGCAACCCCATCATTTAC 1177
QY      288 PhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuLeuValLeuGln 307
Db      1178 TTCTTGTGGGCTCTCTTTAGCAGCGTCGCAAAATAGAGAACTGGAAGCTGTTCTCCAG 1237
QY      308 ArgAlaLeuGlnAspThrProGluValAspGluGlyIleGlyTyrIleuProGlnGluThr 327
Db      1238 AGGGCTCTGAGAGACACGCTGAGGTGATGAAGTGAAGGAGGAGCTTCTCTCAGAAACC 1297
QY      328 LeuGluLeuSerGlySerArgLeuGluGln 337
Db      1298 CTGAGCTGTGAGAGACAAATGAGAGAG 1327

```

```

RESULT 10
E43450
LOCUS    E43450          969 bp    DNA          linear    PAT 31-JAN-2002
DEFINITION Novel protein G-coupled receptor protein and DNA thereof.

```

```

ACCESSION E43450
VERSION   E43450.1 GI:18627716
KEYWORDS  JP 2000166576-A/1.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 969)
AUTHORS   Watanabe,T., Terao,S. and Matsui,H.
TITLE      Novel protein G-coupled receptor protein and DNA thereof
JOURNAL   Patent: JP 2000166576-A 1 20-JUN-2000;
          TAKEDA CHEM IND LTD
          OS      Homo sapiens (human)
          PN      JP 2000166576-A/1
          PD      20-JUN-2000
          PF      30-SEP-1999 JP 1999278116
          PR
          PI      TAKUYA WATANABE, SHIZUKO TERAOKA, HIDEKI MATSUI
          PC      C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00,
          PC      A61P43/00,
          PC      C07K14/705,C07K16/28,C12N1/21,C12N5/10,C12P21/02,C12P21/08, PC
          PC      G01N33/15,
          PC      G01N33/53,G01N33/566,G01N33/577/(C12N1/21,C12R1:19),(C12N5/10, PC
          C12R1:91),
          PC      (C12P21/02,C12R1:19),(C12P21/08,C12R1:91),C12N15/00,A61K37/02,
          PC      C12N5/00,
          PC      (C12N5/00,C12R1:91)
          CC
          FH      Key      Location/Qualifiers
          FT      source      1. 969
                   /organism='Homo sapiens (human)'
FEATURES
source      Location/Qualifiers
             1..969
             /organism='Homo sapiens'
             /mol_type='genomic DNA'
             /db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.:      3,49e-131      Length:      969
Score:          1688.00      Matches:      322
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    95.75%      Indels:      0
DB:             Gaps:      0
US-09-867-570-2 (1-337) x E43450 (1-969)
QY      16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
Db      1 ATGATTCACCAACATCCACAGCTTGGGTACGAACTGACCAATTCACGAGACCTGAGAG 60
QY      36 ThrProCysTyrIleGlnIleThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db      61 ACTCTTGTCTACAGACGACCTGAGCTTCACGAGGCTGACGTGATCGTTTCCCTTGTCTC 120
QY      56 AlaLeuThrGlyAsnAlaValValLeuThrLeuGlyCysArgMetArgAspAsnAla 75
Db      121 GCGCTGACAGAAACCGGCTGTGTGCTCTGCTCTGCGCGCGCATGCCAGAAACGCT 180
QY      76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db      181 GTCTCATCTACATCTCAACCTGATCGCGCGCGCATCTCTCTTACCGGCACAT 240
QY      96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241 ATATGTCGCGGTACGCTCATCAATATCCGCGCATTCCTCAAAATCTCAGTCT 300
QY      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301 GTATGATCACTTCTCTTATAGGCTTAAGCATGTGTAGCGGCATCAGCAGAGCGC 360

```

QY 136 CysLeuSerIleLeuTrpProIleTrpIleHisCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCTGTCATCCTTGTGCCCCATCTGTACCACTGCGCGGCCGACAGAACCTGTATG 420
 QY 156 ValMetCysValIleLeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTrpMetPhe 175
 Db 421 GTCAATGTGTCTCTCTCGGGCCCTGTCTCCGTGCGAGATACCTGTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyValIaAspSerValITrPcysGluThrSerAspPheIleThr 195
 Db 481 TGTACCTTCCTGTGTAGTGTGTCTATCTGTGTGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValIleLeuCysGlySerSerLeuValIleLeuVal 215
 Db 541 ATCCGCTGCTGCTTTTATGTGTGTCTCTGTGGGTCCAGCCTGCTCTGCTGCTGTC 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValIThrIleLeuLeu 235
 Db 601 AGGATTCTCTGTGATCCCGAAGATGCGCTGACAGAGCTGTACGTCACATCCTCTTC 660
 QY 236 ThrValIleValPheLeuLeuCysGlyLeuProPheGlyIleGluIThrAlaLeuPheSer 255
 Db 661 ACAGTGTGCTCTCTCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleHisIleLeuAspTrpIleValIlePheCysHisValIleLeuValSerIlePheLeu 275
 Db 721 AGGATCCACTGTGATTTGAAGTCTTATTTGTGATGTCATCTGATTTCCATTTTCCG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgIle 295
 Db 781 TCCGCTCTTAAACAGAGTGCACACCCCATCATTTACTTCTGCTGCTGCTCTTAAAGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValIleGlnArgAlaLeuGlnAsnTrpProGlu 315
 Db 841 GGTCAAAATAGCCAGAACCTGAAGCTGTCTTCACAGAGCTGTGAGACAGCCTGAG 900
 QY 316 ValAspGluGlyGlyITrPleuProGluIThrLeuGluLeuSerGlySerArgIle 335
 Db 901 GTGATGAAGGTGAGAGGTGCTTCTCTCAGAAACCTGTGAGCTGTGCGGAAGCAGATTG 960
 QY 336 GluGln 337
 Db 961 GAGCAG 966
 RESULT 11
 LOCUS AX498196 969 bp DNA linear PAT 26-SEP-2002
 DEFINITION Sequence 19 from Patent WO0242461.
 ACCESSION AX498196
 VERSION AX498196.1 GI:23343119
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 1
 REFERENCE
 1 Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
 TITLE Endogenous and non-endogenous versions of human g protein-coupled
 JOURNAL receptors
 Patent: WO 0242461-A 19 30-MAY-2002;
 Arena Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..969
 Location/Qualifiers
 1..969
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Novel Sequence"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,49e-131 Length: 969
 Score: 1688.00 Matches: 322
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 95.75% Indels: 0
 DB: 6 Gaps: 0
 US-09-867-570-2 (1-337) x AX498196 (1-969)
 QY 16 MetAspSerThrIleProValIleGluIleGluLeuThrProIleAsnGlyValGluGlu 35
 Db 1 ATGATATTCAACCATCCAGTCTTGGGTACAGAACTGACACCAATCAACGACGTAGAGAG 60
 QY 36 ThrProCysTyrIleGluIThrLeuSerPheThrGlyLeuIThrCysIleValSerLeuVal 55
 Db 61 ACTCTGTGTACAAACAGACCTGAGCTTCACGGGGCTGACGTGATCTGTTCCCTGTGTC 120
 QY 56 AlaLeuIThrGlyAsnAlaValIleValLeuIThrPheLeuGluGlyCysArgMetArgArgAsnAla 75
 Db 121 GCGTGACAGAAAGCGGTTGTCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaIaAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTACATCTCAACCTGGTCCGGCCGACCTTCTCTTACGCGGCACATT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerValIleLeuSerPro 115
 Db 241 ATATGTGCGCCGTACGCTCATCAATCCGCGCATCTCCAAATCCTCAATCTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGACCTTCCCTTACTTTATAGCCTTACAGATGCTGAGCGCATCAGACAGCAGCGC 360
 QY 136 CysLeuSerIleLeuTrpProIleTrpIleHisCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCTGTCATCCTTGTGCCCCATCTGTACCACTGCGCGGCCGACAGAACCTGTATG 420
 QY 156 ValMetCysValIleLeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTrpMetPhe 175
 Db 421 GTCAATGTGTCTCTCTCGGGCCCTGTCTCCGTGCGAGATACCTGTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyValIaAspSerValITrPcysGluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTCTGTGTAGTGTGTCTATCTGTGTGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValIleLeuCysGlySerSerLeuValIleLeuVal 215
 Db 541 ATCCGCTGCTGCTTTTATGTGTGTCTCTGTGGGTCCAGCCTGCTCTGCTGCTGTC 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValIThrIleLeuLeu 235
 Db 601 AGGATTCTCTGTGATCCCGAAGATGCGCTGACAGAGCTGTACGTCACATCCTCTTC 660
 QY 236 ThrValIleValPheLeuLeuCysGlyLeuProPheGlyIleGluIThrAlaLeuPheSer 255
 Db 661 ACAGTGTGCTCTCTCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleHisIleLeuAspTrpIleValIlePheCysHisValIleLeuValSerIlePheLeu 275
 Db 721 AGGATCCACTGTGATTTGAAGTCTTATTTGTGATGTCATCTGATTTCCATTTTCCG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgIle 295
 Db 781 TCCGCTCTTAAACAGAGTGCACACCCCATCATTTACTTCTGCTGCTGCTCTTAAAGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValIleGlnArgAlaLeuGlnAsnTrpProGlu 315
 Db 841 GGTCAAAATAGCCAGAACCTGAAGCTGTCTTCACAGAGCTGTGAGACAGCCTGAG 900
 QY 316 ValAspGluGlyGlyITrPleuProGluIThrLeuGluLeuSerGlySerArgIle 335
 Db 901 GTGATGAAGGTGAGAGGTGCTTCTCTCAGAAACCTGTGAGCTGTGCGGAAGCAGATTG 960
 QY 336 GluGln 337
 Db 961 GAGCAG 966

RESULT 12
 E43451 969 bp DNA linear PAT 31-JAN-2002
 LOCUS Novel protein G-coupled receptor protein and DNA thereof.
 DEFINITION E43451
 E43451.1 GI:18627717
 VERSION JP 2000166576-A/2.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 969)
 Watanabe,T., Terao,S. and Matsui,H.
 Novel protein G-coupled receptor protein and DNA thereof
 Patent: JP 2000166576-A 2 20-JUN-2000;
 TAKEDA CHEM IND LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2000166576-A/2
 PD 20-JUN-2000
 PR 30-SEP-1999 JP 1999278116
 PR TAKUYA WATANABE, SHIZUKO TERAO, HIDEKI MATSUI
 PC C12N15/09, A61K38/00, A61K39/395, A61K45/00, A61K48/00,
 PC A61P43/00,
 PC C07K14/705, C07K16/28, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC
 G01N33/15,
 PC G01N33/53, G01N33/566, G01N33/577// (C12N1/21, C12R1.19), (C12N5/10, PC
 C12R1.91),
 PC (C12P21/02, C12R1.19), (C12P21/08, C12R1.91), C12N5/00, A61K37/02,
 PC C12N5/00,
 PC (C12N5/00, C12R1.91)
 CC
 CC
 CC
 FT source location/Qualifiers
 FT 1..969 location/Qualifiers
 FT /organism='Homo sapiens (human)'.
 FT /db_xref='taxon:9606'
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.08e-131 Length: 969
 Score: 1683.00 Matches: 321
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.69% Mismatches: 0
 Query Match: 95.46% Indels: 0
 DB: Gaps: 0
 US-09-867-570-2 (1-337) x E43451 (1-969)
 QY 16 MetAASPserThrIleProValIleuGluThrGluLeuThrProIleAsnGluYArgGluGlu 35
 DB 1 ATGGATTCAACCATCCCACTGTTGGGTACAGAACTGACCAACCAACGAGCGAGAG 60
 QY 36 ThrProCysTyrIlysglnThrIleuSerPheThrGlyLeuThrCysIleValSerIleVal 55
 DB 61 ACTCCCTTGCACAGAGACCGCTGAGCTTCACGGGGGTGACGTGCATCGTTCCCTTGTG 120
 QY 56 AlALeuThrGlyAsnAlaValIleuThrIleuGluGlyCysArgMetArgAsnAla 75
 DB 121 GCGCTGACAGAAACCGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyIleIle 95
 DB 181 GTCTCCATCTACATCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgIleProIleSerIleLeuSerPro 115
 DB 241 ATATGTTCCGCGTTACGCTCATCAATATCGGCATCCCATCTCCAAATCTCTAGTCT 300

QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 DB 301 GTGATGACCTTTCCTCACTTATAGGCTTAAGATGCTGAGCGGCATCAGACGACGAC 360
 QY 136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
 DB 361 TGCCTGTCATCTCTGTGGCCCACTGTGATCACTGCGCCGCCCAATACCTTCATCG 420
 QY 156 ValMetCysValIleuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
 DB 421 GTATGATGTGTCTGCTGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 DB 481 TGTGACTTCCCTGTTTATGTGTGCTTAATTCGTTGTGTGTAACAGTCAATTTCATTACA 540
 QY 196 IleAlaThrLeuValPheLeuCysValIleuCysGlySerSerIleValIleuVal 215
 DB 541 ATGCGGTGCTGCTGTTTATGTGTGCTTCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 216 ArgIleLeuCysGlySerArgIlysmetProLeuThrArgLeuTyrValIleLeuLeu 235
 DB 601 AGGATTTCTGTGTGATCCCGAAGATGCGCTGACCAAGCTGTACGTGACCATCTCTC 660
 QY 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrAlaLeuPheSer 255
 DB 661 ACAGTGTGCTGCTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleIleAspLeuSerTyrIlyValIleuPheCysHisValIleValIleValIlePheLeu 275
 DB 721 AGGATTCACCTGATGTGAAGAGCTTATTTGTGATGTGATGTGATGTGATGTGATGTG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
 DB 781 TCGGCTCTTAACAGACAGTGCACACCCCATCATTTACTTTCGAGGCTCTTTAGGACAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValIleuValIleuGlnArgAlaLeuGlnAsnThrProGlu 315
 DB 841 CGTCNAATATGAGCAACCTGAAAGCTGTCTCCAGAGGCTCTGACGAGACACCCCTGAG 900
 QY 316 ValAspGluGlyGlyIlyThrLeuProGluGluThrIleuGluLeuSerGlySerArgLeu 335
 DB 901 GTGATGTAAAGTGAAGGTGTGCTTCTTCAGAAACCTGAGACTGTGGAAGACAGATTG 960
 QY 336 GluGln 337
 DB 961 GAGCAG 966
 RESULT 13
 AX923125 969 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 1 from Patent WO03080659.
 DEFINITION AX923125
 AX923125 GI:40216228
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCES
 1 Peri,K.G., Moffett,S. and Abbran,D.
 Methods and compounds for prevention and treatment of elevated
 intraocular pressure and related conditions
 Patent: WO 03080659-A 1 02-OCT-2003;
 THERATECHNOLOGIES INC. (CA)
 JOURNAL
 FEATURES
 source location/Qualifiers
 1..969
 /organism='Homo sapiens'
 /mol_type='unassigned DNA'
 /db_xref='taxon:9606'
 1..969
 /note='unnamed protein product'
 /codon_start=1

```

/protein_id="CAF02233.1"
/db_xref="GI:40216229"
/translation="MDSTIPVGLTELPINGREPTCYKQTLSPFGTLCTIVSLVALTG
DAVAVMLGCRNRNAVSIYIINLVAAPFLFSGHIICSPILINIRHISILSPVM
TPPFIIGLSMLAISTERSLTPMPITWCHRPRLYSIVLCVLAISLRLISLPMF
CDPLFSGADSWCERSDPTITAMLVFLCVLIGSSLVLLVRLICSRKMLPRLVYTI
LLTVLVFLCGLPFGIOWALFSRIHLDKVLFCVHVALVIFPSALNSANPITVFPVG
SFRORONRNLKVLVQRALQDTPVDEGGWLPQETLISGSRLEQ"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9.08e-131 Length: 969
Score: 1683.00 Matches: 321
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.69% Mismatches: 0
Query Match: 95.46% Indels: 0
DB: 6 Gaps: 0

```

US-09-867-570-2 (1-337) x AX923125 (1-969)

```

QY 16 MetAapSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
DB 1 ATGAGATTCAACCATCCAGTCTTGAGGTACAGAACTGACCAATCAACGACGTGAGAG 60
QY 36 ThrProCysTyrGlyGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
DB 61 ACTCTTCTCTACACAGACACCTTGAGCTTACAGGGCTGACGTGCATCGTTCCCTTGT 120
QY 56 AlaLeuThrGlyAsnAlaValLeuThrPleuLeuGlyCysArgMetArgArgAsnAla 75
DB 121 GCGGTGACAGAGACGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
DB 181 GCTTCATCTACATCTCACTCACTGAGCGGCGGAGCTCTCTCTTCTTACGGCCACAT 240
QY 96 IleCysSerProLeuArgPheLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
DB 241 ATATGTTGCGCGTTACGCTCATCATATCCGCAATCCCATCTCCAAATCTCATCTCT 300
QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
DB 301 GTGATGACCTTCCCTACTTATAGGCTTAAGCATCTGACGCGCCATCAAGCCGAGCGCC 360
QY 136 CysLeuSerIleLeuThrProIleTyrPheHisCysArgArgProArgTyrLeuSerSer 155
DB 361 TGCCTGTCATCTGTGGCCCATCTGAGTACCACTGCGCGCCGCGCCAGATACCTGTATCG 420
QY 156 ValMetCysValLeuLeuThrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
DB 421 GTCAATGTGTGCTGCTGCTGCGGCCCTGTCCCTGCTGCGGAGTATCTCGAGTGTGATG 480
QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
DB 481 TGTGCTTCCTGTTAAGTGTGCTGATCTGTTTGGTGAACACGACAGTTTCATTACCA 540
QY 196 IleAlaThrPheValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
DB 541 ATCGGTGGCTGCTGTTTATATGTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
DB 601 AGGATTCTCTGTGATCCCGGAAGATCGCGCTGACCAAGGTGTACGTACCATCTCTCT 660
QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyTyrIleGlnThrPalaLeuPheSer 255
DB 661 ACAGTGTGCTGCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 256 ArgIleHisLeuAspThrPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
DB 721 AGGATTCACCTGATGGTGAAGCTTATATTTGTCAATGTGATCTAGTTTCATTTCCTG 780
QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295

```

```

DB 781 TCCGCTCTTAACAGACAGGCGCAACCCCATCTTACTCTTCGTGGGCTCTCTTAGGAG 840
QY 236 ArgGlnAsnArgGlnAsnLeuLeuValLeuGlnAlaGlnAlaGlnThrProGlu 315
DB 841 CTTCAAAATAGGCAACACTGAACTGTTCTTCCAGAGGGCTTGCAGACACCGCTGAG 900
QY 316 ValAspGluGlyGlyGlyTyrPhePheProGluThrThrLeuGluLeuSerGlySerArgLeu 335
DB 901 GTGATGAAAGTGTGAGGCTGCTCTCTCAAGAAACCTTGAGCTGTCTCGGAAGACGATTG 960
QY 336 GlnGln 337
DB 961 GAGCAG 966

```

```

RESULT 14
AF474987 969 bp DNA linear PRI 16-JUN-2003
LOCUS Homo sapiens G protein-coupled receptor SNSR1 gene, complete cds.
DEFINITION AF474987
ACCESSION AF474987
VERSION AF474987.2 GI:31747894
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

```

```

REFERENCE
AUTHORS Lembo, P.M., Grazzini, E., Groblewski, T., O'Donnell, D., Roy, M.-O.,
Zhang, D., Hoffert, C., Cao, J., Schmidt, R., Pelletier, M., Labarre, M.,
Gosselin, M., Fortin, Y., Banville, D., Shen, S., Strom, P., Payza, K.,
Dray, A., Walker, P., and Ahmad, S.
Proenkephalin A gene products activate a new family of sensory
neuron-specific GPCRs
JOURNAL Nat. Neurosci. 5 (3), 201-209 (2002)
MEDLINE 21853733
PUBMED 11850634

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS Ahmad, S., Banville, D., Fortin, Y., Lembo, P.M. and O'Donnell, D.
TITLE Direct Substitution
JOURNAL Submitted (25-JAN-2002) Molecular Biology, Astrazeneca, 7171
Frederick-Banting, Montreal, Quebec H4S 1Z9, Canada
COMMENT On Jun 16, 2003 this sequence version replaced gi:119338907.
FEATURES
source
1..969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..969
/product="G protein-coupled receptor SNSR1"
1..969
/note="GPCR"
/codon_start=1
/product="G protein-coupled receptor SNSR1"

```

```

CDS
1..969
/note="GPCR"
/product="G protein-coupled receptor SNSR1"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 4.2e-130 Length: 969
Score: 1675.00 Matches: 319
Percent Similarity: 99.69% Conservative: 2
Best Local Similarity: 99.07% Mismatches: 1
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0

```

US-09-867-570-2 (1-337) x AF474987 (1-969)

```

Qy 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
Db 1 ATGATTCACCAACATCCAGTCTGGGTACGAACCTGACACCAATCAACGAGCGAGAG 60
Qy 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db 61 ACTCCCTTGCACAGACGACCTGAGCTTACGGGGGTGACGTGCATCTTCCCTTGC 120
Qy 56 AlaLeuThrGlyAsnAlaValIleuThrPleuGlyCysArgMetArgArgAsnAla 75
Db 121 GCGCTGACAGAAACGCGGTGGCTGCTGGCTCCTGGGCTGCCGACATCCAGGAAGCT 180
Qy 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyHisIle 95
Db 181 GTCTCCATCAACATCCCAACCTGCTGCGCGCCGACCTCTCTTCTTACGGCCACATT 240
Qy 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db 241 ATATGTTCCGCGTTACGCGCTCAATCAATATCCGCAATCTCCAAATCTCAAGTCT 300
Qy 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db 301 GTGATGACCTTCCCTCACTTATAGCCCTAAGCATGCTGAGCGCCATCAGACCGAGCG 360
Qy 136 CysLeuSerIleLeuThrProIleTyrPheCysArgArgProArgTyrLeuSerSer 155
Db 361 TGCCTGCTCCATCTGCGCCCATCTGCTACACGCGCCGACCATCTCTCATCTG 420
Qy 156 ValMetCysValIleuLeuThrPheIleuSerLeuLeuArgSerIleGluTyrPhePhe 175
Db 421 GTATGATGATCTCTGCTGCGCCCTCTCCCTGCTGCGAGATCTCTGAGTGGATGTC 480
Qy 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db 481 TGTGACTCTCTGTTTACGTGGGCAATCTGTTGGTGAAGTGAAGTTCATTATCA 540
Qy 196 IleAlaTyrPheValPheLeuCysValIleuCysGlySerSerLeuValLeuVal 215
Db 541 ATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Qy 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db 601 AGAATCTCTGTGGATCCCGAAGATCCCGCTGACCGAGCTGACGATCTCTCTC 660
Qy 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAlaLeuPheSer 255
Db 661 ACAATGCTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Qy 256 ArgIleHisLeuAspTyrPheValIleuPheCysHisValHisLeuValSerIlePheLeu 275
Db 721 AGGATCCACCTGGATGGAGAGTCTTATTTGTCATGTCATGATCTGATCTGATCTGATCT 780
Qy 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db 781 TCGGCTTTACAGCAGATGCGCAACCCATCAATTAATCTTCTGCGGCTCTCTTACG 840
Qy 296 ArgGlnAsnArgGlnAsnLeuValLeuValIleuGlnArgAlaLeuGlnAsnThrProGlu 315
Db 841 CGTCAAAATAGCGAGAACCTGAGCTGTTCTCCAGAGGCTCTGACGAGACACCCCTGAG 900
Qy 316 ValAspGluGlyGlyTyrPleuProGlnIleuThrLeuGluLeuSerGlySerArgLeu 335
Db 901 GTGATGAAAGTGGGGGTGGCTTCTCTCAAGAAACCTGAGAGCTGTGGGAGAGCAATTTG 960
Qy 336 GluGln 337
Db 961 GAGCAG 966

```

```

RESULT 15
LOCUS AR477199 969 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 4 from patent US 6696257.
ACCESSION AR477199

```

```

VERSION AR477199.1 GI:47234473
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 969)
AUTHORS Ahmed, S., Banville, D., Fortin, Y., Lembo, P., O'Donnell, D. and
Shen, S.-H.
JOURNAL G protein-coupled receptors from the rat and human
FEATURES Patent: US 6696257-A 4 24-FEB-2004.
source Location/Qualifiers
1..969
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,32e-127 Length: 969
Score: 1642.00 Matches: 313
Percent Similarity: 98.76% Conservative: 5
Best Local Similarity: 97.20% Mismatches: 4
Query Match: 93.14% Indels: 0
Gaps: 0
US-09-867-570-2 (1-337) x AR477199 (1-969)
Qy 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
Db 1 ATGATTCACCAACATCCAGTCTGGGTACGAACCTGACACCAATCAACGAGCGAGAG 60
Qy 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db 61 ACTCCCTTGCACAGACGACCTGAGCTTACGGGGGTGACGTGCATCTTCCCTTGC 120
Qy 56 AlaLeuThrGlyAsnAlaValIleuThrPleuGlyCysArgMetArgArgAsnAla 75
Db 121 GCGCTGACAGAAACGCGGTGGCTGCTGGCTCCTGGGCTGCCGACATCCAGGAAGCT 180
Qy 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyHisIle 95
Db 181 GTCTCCATCAACATCCCAACCTGCTGCGCGCCGACCTCTCTTCTTACGGCCACATT 240
Qy 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db 241 ATATGTTCCGCGTTACGCGCTCAATCAATATCCGCAATCTCCAAATCTCAAGTCT 300
Qy 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db 301 GTGATGACCTTCCCTCACTTATAGCCCTAAGCATGCTGAGCGCCATCAGACCGAGCG 360
Qy 136 CysLeuSerIleLeuThrProIleTyrPheCysArgArgProArgTyrLeuSerSer 155
Db 361 TGCCTGCTCCATCTGCGCCCATCTGCTACACGCGCCGACCATCTCTCATCTG 420
Qy 156 ValMetCysValIleuLeuThrPheIleuSerLeuLeuArgSerIleGluTyrPhePhe 175
Db 421 GTATGATGATCTCTGCTGCGCCCTCTCCCTGCTGCGAGATCTCTGAGTGGATGTC 480
Qy 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db 481 TGTGACTCTCTGTTTACGTGGGCAATCTGTTGGTGAAGTGAAGTTCATTATCA 540
Qy 196 IleAlaTyrPheValPheLeuCysValIleuCysGlySerSerLeuValLeuVal 215
Db 541 ATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Qy 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db 601 AGAATCTCTGTGGATCCCGAAGATCCCGCTGACCGAGCTGATCGATCAATCTCTCTC 660
Qy 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAlaLeuPheSer 255
Db 661 ACAATGCTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720

```

QY	256	ArgIleHisLeuAspTrpIysValLeuPheCysHisValHisLeuValSerIlePheLeu	275
Db	721	AGGATCCACCTGGATTGGAAGCTTATTGTGATGTCATCTAGTTTCATTTCCTG	780
QY	276	SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgIn	295
Db	781	TCCGCTCTTAACAGCAGTGCACCAACCCCATCATTTACTTCTTCGTGGGCTCCTTTAGGCAG	840
QY	296	ArgGlnAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGln	315
Db	841	CGTCAAAATAGGCAGAAACCTGAAAGCTGGTTCTCCAAAGGGCTCTGCAGACACGCTGAG	900
QY	316	ValAspGluGlyGlyTyrLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu	335
Db	901	GTGGATGAAGGTGAGGGTGGCTTCTCTCAGGAAACCTGAGACTGTGGGAGACCAATTG	960
QY	336	GluGln	337
Db	961	GAGCAG	966

Search completed: December 2, 2004, 00:13:55
 Job time : 4083 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 1, 2004, 21:28:03 ; Search time 486 Seconds
(without alignments)
3640.031 Million cell updates/sec

Title: US-09-867-570-2
Sequence: 1 MESKSKWVIRLGLSLMDSTI.....EGGGWLPQETLBSGRLEQ 337

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xip
-Q=/cgn2.1/USPTO.spool.p/US09867570/runat.01122004.161122.6327/app.query.fasta_1.519
-DB=Geneseq.23Sep04 -QPM=faastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -INITLS=bits -START=1 -END=1 -MATRIX=biomem2 -TRANS=human40.cdt
-LIST=45 -DOCLIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=Pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09867570 @CGN 1.1 708 @runat.01122004.161122.6327 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	100.0	2618	6	ABK52822 cDNA enco
2	1727.5	98.0	8622	6	ABK52823 Genomic D
3	1722.5	97.7	1369	10	ADC68820 Human GPC
4	1721	97.6	1400	8	AAD33751 Human Mrq
5	1721	97.6	1400	8	ABZ42595 Human G p
6	1721	97.6	1400	12	ADH08534 DNA seque

7	1688	95.7	966	12	AD044603	Ad044603 Human HIT
8	1688	95.7	969	3	AA29811	AA22811 Human G p
9	1688	95.7	969	6	ABT04875	ABT04875 Human G p
10	1688	95.7	969	12	AD030080	Ad030080 Human GPC
11	1688	95.7	1683	10	ADF70583	Adf70583 Orphan re
12	1683	95.5	966	12	AD044605	Ad044605 Human HIT
13	1683	95.5	969	3	AA29812	AA29812 Human G p
14	1683	95.5	969	12	ADF29104	Adf29104 Human GPC
15	1663	94.3	969	3	AAA70342	AAa70342 Human nov
16	1642	93.1	969	2	AA210067	AAa210067 Human dor
17	1598	90.6	969	2	AA210068	AAa210068 Human dor
18	1557	88.3	909	8	ACC44771	ACC44771 Human G p
19	1416	80.3	1369	10	ADC86444	Adc86444 Human GPC
20	1416	80.3	2040	6	AAD33744	Ad33744 Human Mrq
21	1416	80.3	2040	12	ADH08519	Adh08519 DNA seque
22	1416	80.3	2040	12	AD036620	Ad036620 Human mas
23	1405	79.7	969	2	AA210069	AAa210069 Human dor
24	1399	79.4	912	10	ADC12765	Adc12765 Human GPC
25	1397	79.2	997	6	AD116635	Ad116635 Human NOV
26	1397	79.2	997	12	ADN42289	Adn42289 Human CDN
27	1390	78.8	1369	10	ADC86588	Adc86588 Human GPC
28	1388	78.7	1604	6	ABZ42602	Abz42602 Human Mrq
29	1388	78.7	1604	8	ABZ42602	Abz42602 Human Mrq
30	1388	78.7	1604	12	ADH08536	Adh08536 DNA seque
31	1387	78.7	969	3	AA122008	AAa12208 Human G p
32	1387	78.7	969	10	ADC17729	Adc17729 Human TGR
33	1383	78.4	969	4	AAD20943	Ad20943 Human G p
34	1383	78.4	969	5	AA507944	AA507944 Human CDN
35	1383	78.4	969	6	AAD27494	Ad27494 Human G-P
36	1383	78.4	969	6	ABZ42891	Abz42891 Human GPC
37	1383	78.4	969	8	ACC44772	ACC44772 Human G p
38	1383	78.4	969	8	ABZ81334	Abz81334 Human Dor
39	1383	78.4	969	10	AD61647	Ad61647 Human GPC
40	1383	78.4	969	11	ADL96465	Adl96465 Human G p
41	1383	78.4	969	12	AD030078	Ad030078 Human GPC
42	1383	78.4	1116	4	AAH49507	AAh49507 Human GTP
43	1382	78.4	969	2	AAZ10071	AAz10071 Human dor
44	1382	78.4	969	2	AAZ10070	AAz10070 Human dor
45	1367	77.5	969	3	AAA12204	AAa12204 Human G p

ALIGNMENTS

RESULT 1	ABK52822	standard; cDNA; 2618 BP.
ID	ABK52822	
XX	AC	ABK52822;
XX	DT	27-AUG-2002 (first entry)
XX	DE	cDNA encoding human G-protein coupled receptor (GPCR).
XX	KW	Human, G-protein coupled; receptor; GPCR; human protease;
XX	KW	human therapeutic protein; query sequence; search; gene; ss;
XX	KW	sequence database; non-human transgenic animal; gene therapy;
XX	KW	chromosome 3.
XX	OS	Homo sapiens.
XX	FT	
XX	FT	Key
XX	FT	CDS
XX	FT	Location/Qualifiers
XX	FT	/*tag= a
XX	FT	/product= "Human G-protein coupled receptor (GPCR)"
XX	FN	W0200234914-A1.
XX	PD	02-MAY-2002.
XX	PF	10-OCT-2001; 2001WO-US031592.
XX	PR	25-OCT-2000; 2000US-00695045.
XX	PR	31-MAY-2001; 2001US-00867570.

XX (PEKE) PE CORP NY.
 PA
 XX
 PI Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
 XX WPI; 2002-463360/49.
 DR P-PSDB; AAU97598.
 XX Novel isolated G-protein coupled receptor peptide useful for treating
 PT disorder characterized by absence of, in appropriate or unwanted
 PT expression of the receptor protein, and as immunogens to raise
 PT antibodies.
 XX
 PS Claim 4; Fig 1; 75pp; English.
 XX
 CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC peptide. The invention is useful for identifying a modulator of GPCR and
 CC for treating a disease or condition mediated by a human protease. The
 CC invention is also useful as models for the development of human
 CC therapeutics, for identifying therapeutic proteins, as targets for
 CC development of human therapeutic agents, and as query sequence to perform
 CC a search against sequence databases to, for e.g., identify other family
 CC members of related sequences. The vector of the invention is useful for
 CC producing a GPCR protein or peptide, for conducting cell-based assays
 CC involving the GPCR protein or its fragment, for identifying GPCR protein
 CC mutants whose functions are affected, and to produce non-human transgenic
 CC animals. The present nucleic acid sequence represents the human G-protein
 CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
 CC encodes the human G-protein coupled receptor (GPCR) protein of the
 CC invention
 XX
 SQ Sequence 2618 BP; 655 A; 637 C; 546 G; 780 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,876-155 Length: 2618
 Score: 1763.00 Matches: 337
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-867-570-2 (1-337) x ABK52822 (1-2618)

Db 867 TGGCCACTGTGGNACACTGCGCCGCCAGATACCTGTATCATGTGATGTCTCG 926
 QY 161 LeuTPAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPhe 180
 Db 927 CTCGTGGCCCTCTCCCTGTGCGAGATCTCGAGTGAATGTTCTGACTCTCTGTTT 986
 QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
 Db 987 AGTGTGTGATCTCTGTTGGTGTGAAGCTGATTCATTTCATTCATTCGCTGTGTT 1046
 QY 201 PheLeuCysValValLeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGly 220
 Db 1047 TTTTATGTGTGTTCTCTGTGGTTCAGCCGTGCTGTGTGTCAGGATTCCTGTGGA 1106
 QY 221 SerArgIleMetProLeuThrArgLeuValTrpIleLeuLeuThrValLeuValPhe 240
 Db 1107 TCCCGGAAATGCTCCGTGACAGGCTGACGATTCCTTCACAGTCTGTGTTTC 1166
 QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleIleLeuAsp 260
 Db 1167 CTCCTGTGTGCTGCTGCTGCTTGTGATTCAGTGGCCCTGTTTCCAGATTCACCTGAT 1226
 QY 261 TrpIleValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
 Db 1227 TGGAAAGTCTTATTTGTGATGTGATCATGATTTCATTTCCTGTCCGCTTAAACAGC 1286
 QY 281 SerAlaAspProIleIleIleTrpPhePheValGlySerPheAlaArgIleAsnArgGln 300
 Db 1287 AGTGCACACCCATCATTTATCTTCTGTGGGCTCTTTAGGACAGCGTCAAAATAGGAGC 1346
 QY 301 AsnLeuIleLeuValLeuGlnArgAlaLeuGlnAspTrpProGluValAspGluGly 320
 Db 1347 AACCTGAAGCTGTCTCTCCAGAGGCTCTGACAGACACGCTGAGTGGATGAAAGTGA 1406
 QY 321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1407 GGGTGGCTTCTCAGGAAACCTGTGAGGTGTGCGAAGACAGATTGAGACAG 1457

RESULT 2
 ID ABK52823 standard; DNA; 8622 BP.
 AC ABK52823;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Genomic DNA encoding human G-protein coupled receptor (GPCR).
 KW Human; G-protein coupled; receptor; GPCR; human protease;
 KW human therapeutic protein; query sequence; search; gene; ds;
 KW sequence database; non-human transgenic animal; gene therapy;
 KW chromosome 3.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 XX CDS 4300..8496
 FT /*tag= a
 FT /product= "Human G-protein coupled receptor (GPCR)"
 FT exon 4300..4319
 FT /*tag= b
 FT /number= 1
 FT intron 4320..7502
 FT /*tag= c
 FT /number= 1
 FT exon 7503..8496
 FT /*tag= d
 FT /number= 2
 XX
 XX MO200234914-A1.
 XX
 XX 02-MAY-2002.
 XX
 XX

PF 10-OCT-2001; 2001WO-US031592.
 XX
 PR 25-OCT-2000; 2000US-00695045.
 PR 31-MAY-2001; 2001US-00867570.
 XX
 PA (PEKE) PR CORP NY.
 XX
 PI Mei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
 DR WPI; 2002-463360/49.
 DR P-SDB; AAU97598.
 XX
 PT Novel isolated G-protein coupled receptor peptide useful for treating
 PT disorder characterized by absence of, in appropriate or unwanted
 PT expression of the receptor protein, and as immunogens to raise
 PT antibodies.
 PS
 PS Claim 4; Fig 3; 75pp; English.
 XX
 CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC peptide. The invention is useful for identifying a modulator of GPCR and
 CC for treating a disease or condition mediated by a human protease. The
 CC invention is also useful as models for the development of human
 CC therapeutics, for identifying therapeutic proteins, as targets for
 CC development of human therapeutic agents, and as query sequence to perform
 CC a search against sequence databases to, for e.g., identify other family
 CC members of related sequences. The vector of the invention is useful for
 CC producing a GPCR protein or peptide, for conducting cell-based assays
 CC involving the GPCR protein or its fragment, for identifying GPCR protein
 CC mutants whose functions are affected, and to produce non-human transgenic
 CC animals. The present nucleic acid sequence represents the human G-protein
 CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
 CC encodes the human G-protein coupled receptor (GPCR) protein of the
 CC invention
 XX
 SQ Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,04e-151 Length: 8622
 Score: 1727.50 Matches: 332
 Percent Similarity: 96.81% Conservative: 2
 Best Local Similarity: 96.23% Mismatches: 2
 Query Match: 97.99% Indels: 9
 DB: 6 Gaps: 1
 US-09-867-570-2 (1-337) x ABK52823 (1-8622)
 QY 2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12
 DB 7459 CAGAGATCAAAACAGCTGGTATCATCTGTTCTTTCCAGAGGTCATCAGCTGG 7518
 QY 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrgluLeuThrProIleAsnGly 32
 DB 7519 TTTCTGAGCATGAGATTACCATCCAGTCTTGGGTACAGAACGACCATCAAGCGA 7518
 QY 33 ArgGluGluThrProCysTyrLysGlnThrLeuSerPheThrGlyLeuThrCysValleVal 52
 DB 7579 CGTGAGAGAGACTCTTCTCTCAAGCAGACCTGAGCTTCAAGGGGCGAGCTGATCGTT 7638
 QY 53 SerLeuValAlaLeuThrGlyAsnAlaValValLeuThrLeuLeuGlyCysArgMetArg 72
 DB 7639 TCCCTTGTCCCGCGAGAGAAACGGCTGTGTCTGTGGCTCCCGGGCTCCCGCATGGC 7698
 QY 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSer 92
 DB 7699 AGGAGCGCTGCTTCATCTCATCTCACTGCTGCGGCGCACTTCTTCTCTTACG 7758
 QY 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerLysIle 112
 DB 7759 GGGCACATTTATGTTGGCGTTACGGCTCATCATATCCGCCCATCTCCAAATC 7818
 QY 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132

DB 7819 CTCAGTCTGTGATGAGACTTCCCTTATAGGCTTAAGCATGCTGAGCCATCAGC 7878
 QY 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyr 152
 DB 7879 ACCGAGCGTGCCTGCTTCATCTGTGGCCCATCTGTAACACATGCCGCCCCAGATAC 7938
 QY 153 LeuSerSerValMetCysValIleLeuThrPAlaIleSerLeuAsnArgSerIleLeuGlu 172
 DB 7939 CTGTCAATCGGTCAATGTGTGTCTGCTGGGCCCTGTCTCCCTGCGCGAGATCTCTGAG 7998
 QY 173 TrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
 DB 7999 TGGATGTTCTGTATCTTCTCTTGTAGGCGCTGATTTCTTTGGTGTAAACCTCATGAT 8058
 QY 193 PheIleThrIleAlaTrpLeuValPheLeuCysValIleLeuCysGlySerSerLeuVal 212
 DB 8059 TTCATTCATATCCGTCGCTGCTGTTTATTTATGTGTGTCTCTGTGGGTCAACCTGTGC 8118
 QY 213 LeuLeuValArgIleLeuLeuCysGlySerArgLysMetProLeuThrArgLeuTyrValThr 232
 DB 8119 CTGCTGTCAAGATCTCTGTGTGATCCCGGAAAGATCCGCTGACAGGCTGTACGTGACC 8178
 QY 233 IleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAla 252
 DB 8179 ATCTCTCTCAAGTGTGCTGCTTCTCTCTGTGCTGCTGCTTGGCATTCAGTGGGCC 8238
 QY 253 LeuPheSerArgIleHisLeuAspTrpLysValLeuPheCysHisValHisLeuValSer 272
 DB 8239 CTGTTTCCAGATTCACCTGAGTGAAGTGTAAAGTCTTATTTGTGATGATGATCACTTCC 8298
 QY 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySer 292
 DB 8299 ATTTTCTGTGCCCTCTTAAACAGACAGTCCACACCCATCACTTCACTTCTGCGGCTCC 8358
 QY 293 PheArgGlnArgIleAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsp 312
 DB 8359 TTTAGCGAGCTCAAAATAGGCAAGAACCTGAAGCTGTTTCCAGAGGCTCTGCAGGAC 8418
 QY 313 ThrProGluValAspGluGlyGlyTyrPheLeuProGlnGluThrLeuGluLeuSerGly 332
 DB 8419 ACCGCTGAGTGAATGAAGTGAAGGTGCTTCTTCAGAAACCTCGAGCTGTCCGGA 8478
 QY 333 SerArgLeuGluGln 337
 DB 8479 AGCAGATTGAGACAG 8493
 RESULT 3
 AD068820
 ID AD068820 standard; DNA; 1369 BP.
 XX
 AC AD068820;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human GPCR gene SEQ ID NO:1273.
 XX
 KW de; gene; human; GPCR;
 KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN BP1270724-A2.
 XX
 PD 02-JAN-2003.
 XX
 PF 18-JUN-2002; 2002EP-00013517.
 XX
 PR 18-JUN-2001; 2001JP-00246789.
 XX
 PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX
 PI Suwa M, Asai K, Akiyama Y, Aburatani H;

[illegible]

Db	732	TTCAATTCAAACTGGGCTGGCTGGTTTATTATGTCGTGTCCTGTCGGATCCAGCTGGTC	791
Qy	213	LeuLeuValaIaGgIleuCySgISeSaRphYsKePProLeuThrArGleuTyValThr	232
Db	792	CTGCTGGACAGAAATTCCTGTGGAATCCCGAAGATGCCGCTGACCAAGCTGTACGTGACC	851
Qy	233	IleLeuLeuThrValLeuValaPheLeuLeuCySgIleuProPheGylIleGInTPaIa	252
Db	852	ATCTCTCCACAGATGCTGTGCTTCTCCTCTGTGTGGCCCTTTGGCATTCACTGGGCC	911
Qy	253	LeuPheSeSaRgIleHIsleuAspTrpIySValLeuPheCyShIsValHIsleuValSer	272
Db	912	CTGTTTTCACAGATCCACTGTAATGGAAAGTCTTAATTTGTCAATGTCATCTGTGTTCC	971
Qy	273	IlePheLeuSeSaRrAlaLeuAsnSeSeSaRAlaSnProIleIleTyRPhPheValGylSer	292
Db	972	ATTTTCCGTGCTCCGCTCTTAAACAGAGTGCACAACTTCATTTACTTCTGTGGCTCC	1031
Qy	293	PheArGInArGInaSnArGInaSnLeuPheValLeuGInArGAlaLeuGInaSp	312
Db	1032	TTTAGCGAGCTCAAAATAGCCACAACTGTAAGCTGTCTCCAGAGGCTCTGCAAGAC	1091
Qy	313	ThrProGInaIaSpGInGylGylGylTyRLeuProGInGInuThrLeuGInLeuSeSaRyl	332
Db	1092	ACGCTGAGGTGATGAAGGTGAGAGGTGCTTCTCAGAAACCTGTGAGCTGTGGGA	1151
Qy	333	SeSaRgLeuGInGIn 337	
Db	1152	AGCAGATTGAGCAG 1166	
RESULT 4			
ID	AAD33751	standard; DNA; 1400 BP.	
AC	AAD33751;		
DT	01-JUL-2002	(first entry)	
DE	Human MrGX3 (mas-related gene) DNA.		
KM	Human; mas-related gene; G-protein coupled receptor; drg-12 protein;		
XX	receptor; sensory perception; pain; analgesic; MrGX3; gene; ds.		
OS	Homo sapiens.		
Key	Location/Qualifiers		
CDS	332..1300		
FT	/*tag= a		
FT	/product= "Human MrGX3 protein"		
XX	WO200183555-A2.		
XX	08-NOV-2001.		
XX	04-MAY-2001; 2001MO-US014519.		
XX	04-MAY-2000; 2000US-0202027P.		
XX	01-AUG-2000; 2000US-0222344P.		
XX	03-NOV-2000; 2000US-00704707.		
XX	19-APR-2001; 2001US-0285493P.		
XX	(CALY) CALIFORNIA INST OF TECHNOLOGY.		
XX	Anderson DJ, Dong X, Zylka M, Han S, Simon M;		
XX	WPI; 2002-171346/22.		
XX	P-PBDB; AAE21296.		
XX	Isolated polypeptide, MrG, which is a G-protein coupled receptor and an		
XX	isolated polypeptide, drg-12, which is also a receptor, useful for		
XX	identifying agonists or antagonists for treating pain.		
XX	Disclosure; Page 128-129; 185pp; English.		

XX The invention relates to Mrx (mas-related gene) protein, which is a G-protein coupled receptor and drg-12 protein, which is a receptor. The invention is useful for identifying compounds that bind to it, especially agonists or antagonists. Administration of an agent (e.g. the identified agent) that increases the expression of Mrx in a mammal may be used for treating impaired sensory perception in a mammal, especially pain. The agent may also be useful for treating impaired sensory perception in a mammal. The present sequence is human Mrx3 DNA

Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.09e-151	Length:	1400
Score:	1721.00	Matches:	330
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	97.62%	Indels:	0
DB:	6	Gaps:	0

US-09-867-570-2 (1-337) x AAD33751 (1-1400)

Qy 6 SerTryValIleAArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 Db 302 TCCAGGGTCAACAGACTGGGGTTCTGAGCATGATTCACACATCCAGTCTTGGAACA 361
 Qy 26 GluLeuThrProIleAangIyArgGluThrProCysTyrIlyGlnThrLeuSerPhe 45
 Db 362 GAACGTACACCAATCAACGAGCGTGAAGAGACTCTTGCTCAAGACAGACCTGAGCTTC 421
 Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAmaAlaValLeuThr 65
 Db 422 ACCGGGCTGACGGTCACTGTTCCCTTGTCCGGTGAAGAAACGGGTTGCTCTCG 481
 Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleAsnLeuValAla 85
 Db 482 CTCTGGGCTGCCCATGCGAGAACGCTGTCTTCATCTACATCTTCACCTGTCGCG 541
 Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
 Db 542 GCGACTTCTCTCTCTTAGCGGCCACATATATGTTCCGCGTTACGCTCATCATATC 601
 Qy 106 ArgHisProIleSerIyIleLeuSerProValMetThrPheProIyrrPheIleGlyLeu 125
 Db 602 CGCCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCCCTACTTATAGGCTTA 661
 Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleTyr 145
 Db 662 AGCATGCTGAGCGCATCAGCACCGAGCGCTGCTTCATCTCTGTGCGCCATCTGATC 721
 Qy 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuThrPalaLeuSer 165
 Db 722 CACTGCGCGCGCCCAATACCTGTCACTGCTGATGATGTCCTGCTGCGCTGCGCTG 781
 Qy 166 LeuLeuArgSerIleLeuGluThrPheCysAspPheLeuPheSerGlyAlaAspSer 185
 Db 782 CTGCTGGAGATACCTGAGATGATGTTCTGTACTTCCGTTAGTGCGCTGATTCCT 841
 Qy 186 ValTyrCysGluThrSerAspPheIleThrIleAlaThrLeuValPheLeuCysValVal 205
 Db 842 GTTTGGTGAACGTACGATTTTCAATACCGCGTGGCTGTTTATATGATGCTT 901
 Qy 206 LeuCysGlySerSerLeuValLeuValArgIleLeuCysGlySerArglyMetPro 225
 Db 902 CTCTGTTGGTCAAGCTGCTCTCTGATGAGATTCCTGTGATCCCGAAGATGCGG 961
 Qy 226 LeuThrArgLeuThrValThrIleLeuLeuThrValLeuValPheLeuCysGlyLeu 245
 Db 962 CTGACCGAGGCTGACGATCATCTCTCAAGATGCTGCTCTCTCTCTGAGGCTG 1021
 Qy 246 ProPheGlyIleGlnTyrPalaLeuPheSerArgIleHisLeuAspTyrPalaLeuPhe 265
 Db 1022 CCTTTGGCATTCAGTGGCGCTCTTTTCCAGATCCACCTGATTTGGAAGTCTTATTT 1081

Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 Db 1082 TGTCATGTGATCATGATTTTCATTTCTCTGCTCCCTTAAACAGCATGCGCAACCCATC 1141
 Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuValLeuVal 305
 Db 1142 ATTACTTCTTGTGGCTCTCTTTAGCAGCGTCAAAATAGGAGAACTGAACTGCTT 1201
 Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAlaAspGluGlyIlyTyrLeuProGln 325
 Db 1202 CTCAGAGGCTCTGACAGACAGCCCTGAGGTGATTAAGTGGAGGTGCTTCTCAG 1261
 Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1262 GAAACCTGAGGCTGTGCGGAGACAGATTGAGACAG 1297
 RESULT 5
 AB242595
 ID AB242595 standard; DNA, 1400 BP.
 AC AB242595;
 XX
 XX 04-MAR-2003 (first entry)
 DE Human G protein-coupled receptor Mrx3 nucleotide SEQ ID NO:673.
 XX
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiovascular; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer; gene; ds.
 KM
 OS Homo sapiens.
 XX
 XX W0200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI: 2003-046718/04.
 DR P-PSDB; ABP81750.
 DR
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 XX Disclosure; Fig 1; 523pp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for

CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related diseases, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, cirrhosis, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention

XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.09e-151	Length:	1400
Score:	1721.00	Matches:	330
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	97.62%	Indels:	0
DB:	8	Gaps:	0

US-09-867-570-2 (1-337) x AB242595 (1-1400)

QY 6 SerTPValIleaygLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 DB 302 TCCAGGGTCACACGAGCTGGGTTTCTGACATGATTCACCACTCCAGCTTGGGTAC 361
 QY 26 GIUleuThrProIleasnIyArngIugIuThrProCyetrIyngIlnhrLeuSerpe 45
 DB 362 GAACAGACCAATCAACGAGCGTAGAGACTCTTGCTACAGAGACCTGAGCTTC 421
 QY 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTrp 65
 DB 422 ACGGGGCTGACGTGATCGTTTCCCTGTCGGCTACAGAGAACCGGTTGCTCG 481
 QY 66 LeuLeuGlyCyArngMetArngAsnAlaValSerIleTyrlleLeuAsnLeuValAla 85
 DB 482 CTCCTGGGCTGCGCATGCGAGAGACGCTGCTCCATCATCACTCAACCTGTCGC 541
 QY 86 AlaAspPheLeuPheLeuSerGlyHsiIleLeuCySerProLeuArgLeuIleAsnIle 105
 DB 542 GCCGCTTCTCTTCTTACCGGCCACATTATATGTCGGCTTACGCCCTCATCATATC 601
 QY 106 ArgHsiProIleSerIySileLeuSerProValMetThrPheProTyrlPheIleGlyLeu 125
 DB 602 CGCCATCCCATCTCCAAATTCCTCATGCTGATATACCTTTCCCTACTTTATAGCCCTA 661
 QY 126 SerMetLeuSerAlaIleSerThrGlyIuArngCyLeuSerIleLeuTrpProIleTyrl 145
 DB 662 AGCAGCTGAGCGCCATCAGCACCGAGCGCTGCTTCCATCTTGCGCCCATCTGTATC 721
 QY 146 HsiCyArngArngProArngTyrlLeuSerSerValMetCyValLeuLeuTrpAlaLeuSer 165
 DB 722 CACTCCCGCGCCCGCATACCTGTCATGCTATGTCGTCTGCTGCGCGCTGCTC 781
 QY 166 LeuLeuArngSerIleLeuGlyTrpMetPheCyAspPheLeuPheSerGlyAlaAspSer 185
 DB 782 CTGCTGCGGAGATCTCTGAGTGGAGTCTTGACTCTCTGTTAGTGTGCTGATTC 841
 QY 186 ValTrpCyGlyIuThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCyValVal 205
 DB 842 GTTGTGTGGAACGTCAGATTTCATTACATGCGTGGGTGTTTATATGTGGT 901
 QY 206 LeuCyGlySerSerIleuValLeuValArgIleLeuCyGlySerArgIySmetPro 225
 DB 902 CTCGTGGGTCCAGCTGTGCTGCTGCTGAGGATTCCTGTGTGATCCCGAAGAGCGC 961
 QY 226 LeuThrArgLeuTyrlValThrIleLeuLeuThrValLeuValPheLeuLeuCyGlyLeu 245

DB 962 CTGACAGGCTTACGTGACCATCTCTCCACAGGCTGCTTCTCTCTGTGGCCTG 1021
 QY 246 PropheGlyIleGlnTrpAlaLeuPheSerArgIleHsiLeuAspTrpIyValLeuPhe 265
 DB 1022 CCGTTGGATTCAGTGGGCCCTGTGTTTCCAGAGTCCACCTGGATTGGAAGTCTTATTT 1081
 QY 266 CyHsiValHsiLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 DB 1082 TGTCTATGTCATCTAGTTTCCATTTTCTGTGCGCTTTTACAGAGAGGCCAACCCATC 1141
 QY 286 IleTyrlPhePheValGlySerPheArngIuArngIuAsnArngIuAsnLeuIyLeuVal 305
 DB 1142 ATTACTTCTTGTGGCTCTCTTTAGGCAAGCTCAAAATAGGCAAGACCTGAAGCTGTT 1201
 QY 306 LeuGlnArngAlaLeuGlnAspThrProGluValAspGlyIyGlyIyTrpLeuProGln 325
 DB 1202 CTCAGAGGGCTCTCAGACACGCTGAGGTGATGAGGTGAGGTGGCTTCTGAG 1261
 QY 326 GIUThrLeuGlyLeuSerGlySerArngLeuGlyGln 337
 DB 1262 GAACCTTGAGCTGTGGGAGACGATTGGAGCAG 1297
 RESULT 6
 ADH08534
 ID ADH08534 standard; DNA; 1400 BP.
 AC ADH08534;
 XX 25-MAR-2004 (first entry)
 DT XX
 DE DNA sequence MrgX3.
 XX mas-related gene D; MrgD; Analgesic; Vulnerary; Ophthalmological;
 KM sensory perception; Glaucoma; Mrg; ds.
 OS Mus musculus.
 PN W02004003133-A1.
 XX 08-JUN-2004.
 PD PF
 XX 13-MAY-2003; 2003WO-US015004.
 PR 26-JUN-2002; 2002US-00183116.
 PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
 PI WPI; 2004-083025/08.
 DR P-PSDB; ADH08535.
 PT New mas-related gene D polypeptides, useful as therapeutics or in
 PT identifying agonists or antagonists that alter pain perception in a
 PT mammal for treating impaired sensory perception, e.g. chronic intractable
 PT pain or neuropathic pain.
 XX disclosure; SEQ ID NO 30; 220pp; English.
 XX The present invention relates to an isolated mas-related gene D (MrgD)
 CC polypeptide. The MrgD polypeptides are useful as therapeutics or for
 CC identifying compounds, i.e. agonists or antagonists, that alter pain
 CC perception in a mammal. The compounds are useful for treating impaired
 CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
 CC promoting wound healing, restoring normal sensitivity following injury,
 CC or treating ocular conditions, particularly those associated with
 CC pressure such as glaucoma. The Mrg genes or proteins may be used as
 CC molecular probes for the detection of cells or tissues related to or
 CC involved with sensory perception. The present sequence represents a MrgA
 CC (Mrg subfamily) encoding sequence.
 XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,09e-151 Length: 1400
 Score: 1721.00 Matches: 330
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 97.62% Indels: 0
 DB: 12 Gaps: 0

US-09-867-570-2 (1-337) x ADH08534 (1-1400)

QY 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 DB 302 TCACGGGCTCACCACTGAGGCTTCTAGCATGATTCACATCCAGCTTGGGTACA 361
 QY 26 GluLeuThrProIleAenGlyArgGluGlyThrProCysTrpGlyGlnThrLeuSerPhe 45
 DB 362 GAACGTACACCAATCAACGAGCGTGAAGAGCTCTTGCTACAGCAGACCTAGCTTC 421
 QY 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAenAlaValValLeuTrp 65
 DB 422 ACGGGGCTGACGTGATCGTTCCCTGTCGCGCTGACAGAAACGGGTTGCTCTGG 481
 QY 66 LeuLeuGlyCysArgMetArgArgAenAlaValSerIleTrpIleAenLeuValAla 85
 DB 482 CTCTGGGCTGCGCAGCAGCGAGCAGCTGTCTCAATCAATCCATCCATGCTGCG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyIleIleCysSerProLeuArgLeuIleAenIle 105
 DB 542 GCGGACTTCTCTTCTTAGGCGCACATTAATGTTGCGCGTACCGCTCATCAATATC 601
 QY 106 ArgHisProIleSerIleValLeuSerProValMetThrPheProIleGlyLeu 125
 DB 602 CGCATCCCATCTCCAAATCTCAGTCTGATGATGATCTTCCCTTATATAGGCTTA 661
 QY 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTrp 145
 DB 662 AGCAGCTGAGCGCAGCAGCAGCAGCGCTGCTGCATCCCTGAGCCCATCTGGTAC 721
 QY 146 HisCysArgArgProArgTrpLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer 165
 DB 722 CACTGCGCGCGCCCAATACCTGTCTCATGCTCATGCTGCTGCGCTGCGCTGCTG 781
 QY 166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
 DB 782 CTCTGCGGAGTATCTCGAGTGTGATGTCTGATCTTCCGTTAGTGGGCGATTCCT 841
 QY 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
 DB 842 GTTGGGTGAACGTCAGATTCATTACATCCGCGCTGCTTTTATATGTGTGTT 901
 QY 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
 DB 902 CTCTGCGGCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
 QY 226 LeuThrArgLeuGlyValIleThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
 DB 962 CTGACCCAGGCTGATCGTACGATCTCTCTCAAGTGTGCTGCTCTCTCTGCGCTG 1021
 QY 246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisIleAspTrpIleValLeuPhe 265
 DB 1022 CCGCTTGGCATTCAGTGGCGCTGCTTCCAGATCCAGCTGAGTGGAAAGCTTATTT 1081
 QY 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 DB 1082 TGCATGTGATCATAGTTTCATTTCTCTGCTGCTTAAACAGCATGCAACCCATC 1141
 QY 286 IleTrpPhePheValGlySerPheArgGlnArgGlnAenArgIleAenLeuValLeuVal 305
 DB 1142 ATTACTTCTTCTGCGCTCTCTTAGCGAGCTCAAAATAGGCGAAGCTGAGCTGTT 1201
 QY 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTrpLeuProGln 325

DB 1202 CTCAGAGGGCTCTGCAAGACCCCTGAGGTGATGAGGTGAGGGTCTTCTCAG 1261
 QY 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 DB 1262 GAACCTGGAGCTGTGCGGAGAGCATTTGGAGCAG 1297
 RESULT 7
 ADO44603
 ID ADO44603 standard; DNA; 966 BP.
 AC ADO44603;
 XX 29-JUL-2004 (first entry)
 DT 29-JUL-2004 (first entry)
 DE Human H177213 protein encoding DNA.
 XX H177213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
 KW cytosolic; nephrotropic; antiinflammatory; dermatological; analgesic;
 KW vulnerrary; neuroprotective; human; gene; de.
 OS Homo sapiens.
 XX
 OS
 PH Key Location/Qualifiers
 FT CDS 1..966
 FT /*tag= a
 FT /product= "H177213"
 PN MO204039972-A1.
 XX
 PD 13-MAY-2004.
 XX
 XX 28-OCT-2003; 2003MO-JP013781.
 PF 29-OCT-2002; 2002JP-00314141.
 XX
 XX (TRAKE) TAKEDA CHEM IND LTD.
 PA Kaisho Y, Watanabe T, Yasuhara Y, Mori I, Taketomi S;
 PI WPI; 2004-376191/35.
 DR P-PSDB; ADO44602.
 XX
 DR H177213 protein, encoded DNA and transgenic animals for clarifying
 PT pathological mechanism, developing therapeutic methods and screening
 PT preventives or remedies for related diseases e.g. cataract, cancer, and
 PT dermatitis.
 XX
 PS Disclosure; SEQ ID NO 2; 161pp; Japanese.
 XX
 XX The invention relates to a non-human mammal that carries a DNA integrated
 CC with a foreign H177213 or its mutant gene, or a part of it. The non-human
 CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
 CC onset, transient skin rash and proliferation-promoting activity. The
 CC foreign H177213 gene is a gene that encodes a G protein-coupled receptor
 CC (GPCR) protein H177213. The protein, its encoded DNA and constructed
 CC transgenic animals are useful for clarifying pathological mechanism,
 CC developing therapeutic methods and screening preventives or remedies for
 CC related diseases e.g. cataract, cancer, and dermatitis. The present
 CC sequence represents a DNA encoding a human H177213 protein.
 XX
 SQ Sequence 966 BP; 170 A; 280 C; 249 G; 267 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.31e-149 Length: 966
 Score: 1698.00 Matches: 322
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.75% Indels: 0
 DB: 12 Gaps: 0
 US-09-867-570-2 (1-337) x ADO44603 (1-966)
 QY 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAenGlyArgGluGln 35

```

Db      1 ATGATTACACATCCAGTCTTGGGTACAGAACTGACACCATCAACGAGCGTGGAG 60
Qy      36 ThrProCysTyrTyrlsGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db      61 ACTCTTCTCTACACAGACAGCCCTGAGCTTCACGGGGCTGACGCGATCTCTTCTCTTGGTC 120
Qy      56 AlaLeuThrGlyAsnAlaValValLeuThrPheLeuGlyCysArgMetLeuArgAsnAla 75
Db      121 GCGGTACAGGAAAGCGGTGTGTGCTCTGCTCTGGCTGGCTGGCCAGTGGCGAGAAAGCT 180
Qy      76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db      181 GTCTCATCTACATCCTCAACCTGGTGGGCGGAGCTTCCTCTCTCTTGGCGGCGACATT 240
Qy      96 IleCysSerProLeuThrLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241 ATATGTTGGCCCTTACGCTCATCATCATATCCGCGCATCTCCAAATCTCAAGTCTCT 300
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301 GTGATGACCTTTCTCTACTTTATAGGCTTAGCATGCTAGCGCCATCAGCACCGAGCGC 360
Qy      136 CysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyrLeuSerSer 155
Db      361 TGCGTGTCCATCTGTGGCCCATCTGTGACACTGCGCGCGCCGACATGCTGTCATCG 420
Qy      156 ValMetCysValLeuLeuThrPalaLeuSerLeuLeuArgSerIleLeuGluThrPhe 175
Db      421 GTCAAGTGTGTCTGTGTGGGCGCTGTCTCTGCTGCGAGTACTGTGAGTGAAGTTC 480
Qy      176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db      481 TGTGACTTCTCTTTAGTGTGTGTGATTTCTGTTGGTGTGAACGTCAGATTTTCATTACA 540
Qy      196 IleAlaTyrPheValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
Db      541 ATCGGTGGCTGGTTTTTTATGTGTGTGTCTGTGTGGGTCCAGCGTGGCTGTGGTGC 600
Qy      216 ArgIleLeuCysGlySerArgTyrMetProLeuThrArgLeuTyrValThrIleLeu 235
Db      601 AGGATTTCTGTGGATCCCGAAGATGCGCTGACCGATGACGACATCTCTCTC 660
Qy      236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrPalaLeuPheSer 255
Db      661 ACAGTGTGTGTCTCTCTCTGTGTGGCTGTCTTGGCATTCAGTGGGCTCTGTTTCC 720
Qy      256 ArgIleHisLeuAspTyrPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db      721 AGGATCCACCTGGATTGAAAGCTTATTTGTGCATGTGCATCTCAATTTCTCTG 780
Qy      276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db      781 TCCGCTCTTAAACGACGTCGCAACCCCATCATTTACTCTTCTGTGGGCTCTTTAGGAG 840
Qy      296 ArgGlnAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsnThrProGlu 315
Db      841 CGTCAAAATAGGCAAGAACTGAAAGCTGTCTCTCAAGAGGCTCTGACAGGACGCGCTGAG 900
Qy      316 ValAspGluGlyGlyTyrPheLeuProGlnIleThrLeuGluLeuSerGlySerArgLeu 335
Db      901 GTGGATGAAAGGTGAGGTGTGCTTCTCTCAGAAACCTGTGAGCTGTGGGAGACAGATTG 960
Qy      336 GluGln 337
Db      961 GAGCAG 966

```

RESULT 8
 AAA29811
 ID AAA29811 standard; cDNA; 969 BP.
 AC AAA29811;
 XX

```

DT      18-AUG-2000 (first entry)
XX      DE
XX      Human G protein-coupled receptor hH17T213 encoding cDNA SEQ ID NO:3.
XX      DE
XX      Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
XX      KW genetic disease; cellular function regulation; ss.
XX      OS Homo sapiens.
XX      FH Key Location/Qualifiers
XX      FT 1..969
XX      FT CDS /tag= a
XX      FT /product= "G protein-coupled receptor"
XX      PN MO200020455-A1.
XX      PD 13-APR-2000.
XX      PF 30-SEP-1999; 99WO-JP005366.
XX      PR 01-OCT-1998; 98JP-00279535.
XX      PA (TAKE ) TAKEDA CHEM IND LTD.
XX      PI Watanabe T, Terao Y, Matsui H;
XX      PI WPI; 2000-303747/26.
XX      DR P-PSDB; AA90761.
XX      PT Human-derived G protein-coupled protein and encoding nucleic acid, useful
XX      PT e.g. in determining ligands and treatment of diseases associated with
XX      PT dysfunction of the protein.
XX      PS Claim 6; Page 93-94; 97pp; Japanese.
XX      CC The present sequence encodes a human-derived G protein-coupled protein
XX      CC designated hH17T213, which is isolated from the human hippocampus. The G
XX      CC protein-coupled receptor can be used for preventing, treating and
XX      CC diagnosing genetic diseases associated with G protein-coupled protein,
XX      CC and for regulating cellular functions. The protein can be used to prevent
XX      CC and treat disorders associated with G protein-coupled protein gene
XX      CC dysfunction. It can also be used to identify G protein-coupled protein
XX      CC ligands and generating antibodies and antisera against the protein. It is
XX      CC also useful in constructing recombinant receptor protein expression
XX      CC systems, developing receptor-binding assay systems and screening drug
XX      CC candidates, and can be used as a probe in the genetic diagnosis of G
XX      CC protein-coupled protein disorders
XX      SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
XX      SO

```

Alignment Scores:
 Prod. No.: 8,34e-149 Length: 969
 Score: 1688.00 Matches: 322
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.75% Indels: 0
 DB: 3 Gaps: 0

US-09-867-570-2 (1-337) x AAA29811 (1-969)

```

Qy      16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
Db      1 ATGATTACACATCCAGTCTTGGGTACAGAACTGACACCATCAACGAGCGTGGAG 60
Qy      36 ThrProCysTyrTyrlsGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db      61 ACTCTTCTCTACACAGACAGCCCTGAGCTTCACGGGGCTGACGCGATCTCTTCTGTGTC 120
Qy      56 AlaLeuThrGlyAsnAlaValValLeuThrPheLeuGlyCysArgMetLeuArgAsnAla 75
Db      121 GCGGTACAGGAAAGCGGTGTGTGCTCTGCTCTGGCTGGCTGGCCAGTGGCGAGAAAGCT 180
Qy      76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95

```

Db 181 GTCTCCATCTACATCTCACTGCTGCGGCGCACTTCTCTTCTTACGCGCACTT 240
 QY 96 ILeCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATATGTTCCGCGTACCGCTCATCATATCCGATCCCATCTCCAAATCTCAGTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTATGACCTTTCCCTACTTATAGGCGCTAAGCATGCTGACGCGCATCAGACGACGCG 360
 QY 136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCCTGCTCCATCTCTGAGCCCATCTGTAACCATGCGCGCGCCCAAGTACTCTCATCG 420
 QY 156 ValMetCysValLeuLeuTyrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
 Db 421 GTCATGTGTGCTGCTGCTGCGGCTGTCTGCTGCGGAGTATCCTGAGTGTGATGTT 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTCTGTTAGTGTGCTGATTCGTGTGTGTGTAACGTCAGATTTCATTACA 540
 QY 196 ILeAlaThrLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATGCGCTGCTGCTGCTTCTTATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 216 ArgIleLeuCysGlySerArgIleCysMetProLeuThrArgLeuTyrValThrIleLeu 235
 Db 601 AGGATTCCTCTGTGATCCCGAAGATGCGCTGACGAGCGCTGACGATTCCTCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrPalaLeuPheSer 255
 Db 661 ACAGTGTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleHisLeuAspTyrPylsValLeuPheCysHisValHisLeuValSerIlePheLeu 275
 Db 721 AGGATTCACCTGATTTGGAAGTCTTATTTGATGATGATGATGATGATGATGATGATGAT 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
 Db 781 TCCGCTCTTAAACAGCAGTGCACCCCATCATTTACTTCTTCTGAGGCTCTTAAAGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsnThrProGln 315
 Db 841 CGTCAAATAGGAGCAACCTGAGCTGTTCTCCAGAGGCTCTGAGGACACCGCTGAG 900
 QY 316 ValAspGluGlyGlyTyrPheProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGATGAGAGTGAAGGTGCTCTCTCAAGAAACCTGAGGCTGTGAGGAGCAGATTG 960
 QY 336 GluGln 337
 Db 961 GAGCAG 966
 RESULT 9
 ID ABT04875 standard; cDNA; 969 BP.
 AC ABT04875;
 XX
 DT 11-OCT-2002 (first entry)
 XX Human G protein coupled receptor hrUP37 coding sequence.
 DE Human; G-protein coupled receptor; GPCR; hrUP28; hrUP29; hrUP30; hrUP31;
 KM hrUP32; hrUP33; hrUP34; hrUP35; hrUP36; hrUP37; gene; ss.
 XX Homo sapiens.
 OS
 XX MO200242461-A2.
 PN
 XX 30-MAY-2002.
 PD

XX 26-NOV-2001; 2001WO-US044386.
 PE 27-NOV-2000; 2000US-0253404P.
 PR 12-DEC-2000; 2000US-0253566P.
 PR 20-FEB-2001; 2001US-0270266P.
 PR 20-FEB-2001; 2001US-0270286P.
 PR 06-APR-2001; 2001US-0282032P.
 PR 06-APR-2001; 2001US-0282356P.
 PR 06-APR-2001; 2001US-0282358P.
 PR 06-APR-2001; 2001US-0282365P.
 PR 14-MAY-2001; 2001US-0290917P.
 PR 31-JUL-2001; 2001US-0309208P.
 XX
 PA (AREN-) ARENA PHARM INC.
 PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
 XX WPI; 2002-566565/60.
 XX P-PSDB; ABU04077.
 DR
 XX Novel endogenous and non-endogenous versions of G protein-coupled
 PT receptor useful for identification of candidate compounds as receptor
 PT agonists or antagonists for use as therapeutic agents.
 XX
 PS Claim 39; Page 74; 84p; English.
 CC The present invention provides the protein and coding sequences of
 CC several human G-protein coupled receptors (GPCRs). These can be used in
 CC the identification of candidate compounds as receptor agonists or inverse
 CC agonists having applicability as therapeutic agents. The present sequence
 CC is a GPCR coding sequence of the invention
 XX
 SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8,34e-149 Length: 969
 Score: 1688.00 Matches: 322
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.75% Indels: 0
 DB: Gaps: 0
 US-09-867-570-2 (1-337) x ABT04875 (1-969)
 QY 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 Db 1 ATGATTCACACCATCCCAAGTCTTGGTACAGAACTGACACCATCAACGACGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysValIleValSerLeuVal 55
 Db 61 ACTCTTGCTAACAAGCAGACCTGAGCTTCAACGAGGCTGACGAGTGCATCGTTTCTCTGTC 120
 QY 56 AlaLeuThrGlyAsnAlaValValLeuThrLeuGluGlyCysArgMetArgAspAla 75
 Db 121 GCGCTGACAGAAACGCGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTACATCTCCCAACCTGCTGCGCGGCGCATCTCTCTCTTCAACGCGCACTT 240
 QY 96 ILeCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATATGTTCCGCGTACCGCTCATCATATCCGATCCCATCTCCAAATCTCAGTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTATGACCTTTCCCTACTTATAGGCGCTAAGCATGCTGACGCGCATCAGACGACGCG 360
 QY 136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCCTGCTCCATCTCTGAGCCCATCTGTAACCATGCGCGCGCCCAAGTACTCTCATCG 420

Db	121	GGCGTGAACAGAAACGGCGTTGGCTGTGAGCTCTCGGCGTGGCCGAGTGGCAGAAAGCT	180
Qy	76	ValSerIleTyrIleLeuAsnLeuValAlaIAspPheLeuPheLeuSerGlyHisIle	95
Db	181	GGTTCACATCTACATCTCTCAACCTGGTGGCGGCGACCTTCCTCTTGGCGGCCACATT	240
Qy	96	IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleValIleuSerPro	115
Db	241	ATATGTTGGCCGATTAGCCCTCAATCAATATCCGACATCCATCTCCAAATCTTCAGTCTC	300
Qy	116	ValMetCtnrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg	135
Db	301	GGATGACCTTTCCCTACTTATAGCCCTTAGATCTCTAGGCCCATACAGACCGAGCCG	360
Qy	136	CysLeuSerIleLeuTTPProIleTTPYrHisCysArgArgProArgTyrLeuSerSer	155
Db	361	TGCCCTGCATCTGTGGGCCCATCTGGTACCACTGCCGCCGCCAGATACCTGATCG	420
Qy	156	ValMetCysValLeuLeuTTPAlaLeuSerIleLeuArgSerIleLeuGluTTPMetPhe	175
Db	421	GCATGTTGTGCTCTGTGGGCCCTGTGCCCTGTGGCGGAGTACCTGGAGTGAATGTC	480
Qy	176	CysAspPheLeuPheSerGlyAlaAspSerValTTPCysGluThrSerAspPheIleThr	195
Db	481	TGTGACTTCCTGTTTATAGTGTGTCTATCTGTTGGTGTGAACGTCAGATTCATTACA	540
Qy	196	IleAlaTTPLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal	215
Db	541	ATCGGTTGGCTGTGTTTTTTATGTGTGTCTCTGTGGGTCAGCCCTGGTCTGTGTC	600
Qy	216	ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu	235
Db	601	AGGATTCTCTGTGGATCCCGGAAGATCCGCGTGAACAGGCTGTACGTGACCATCTCTC	660
Qy	236	ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTTPAlaLeuPheSer	255
Db	661	ACAGTGTGCTCTCTCTCTCTGTGGCCCTGTGGCATTAAGTGGCCCTGTGTTCC	720
Qy	256	ArgIleHisIleLeuAspTTPLeuValLeuPheCysHisValHisIleValSerIlePheLeu	275
Db	721	AGGATCCACCTGGATTGGAAAGCTTATTTGTGCATGTGATCTAATTTCATTTCCTG	780
Qy	276	SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln	295
Db	781	TCCGCTCTTAAACAGCAGTGCACAACTCATATTACTTCTTCGTGGGCTCTTTAGGCA	840
Qy	296	ArgGlnAsnArgGlnAsnLeuIleValIleGlnArgAlaIleGlnIleAspThrProGln	315
Db	841	CGTCAAAATTAAGCAGAACTGAAGCTGGTCTTCACAGAGGCTCTTGACAGCAGCGCTGAG	900
Qy	316	ValAspGluGlyGlyTyrPheLeuProGlnGluThrLeuGluIleuSerGlySerArgLeu	335
Db	901	GTGGATTGAAGTGGAGGCGGTCTCTCTCAAGAAACCTTGAGCTGTGCGGAACAATTG	960
Qy	336	GluGln 337	
Db	961	GAGCAG 966	
RESULT 11			
ADF70583			
ID	ADF70583	standard, DNA, 1683 BP.	
XX	ADF70583;		
XX	AC		
XX	DT	12-FEB-2004 (first entry)	
DE		Orphan receptor ligand-related human protein gene SeqID206.	
XX		ligand; orphan receptor fusion protein; fluorescent protein;	
KW		cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;	
KW		GFPuv; Enhanced GFP; EGFP; human; gene; de.	
XX		Homo sapiens.	
OS			

[illegible]

QY 156 ValMetCysValLeuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluThrPhePhe 175
 Db 421 GTCATGTGTCTCTGCTGAGCCCTGCTCCCTGCTGCGAGATCTCTGAGTGAAGTTC 480
 QY 176 CysAspPheLeuPheSerGlyValIAspSerValITPcYsgluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTCTGTAGTGTCTGATTTCTGTTGGTGTGAACGTCGATTTTCATTACA 540
 QY 196 IleAlaTPLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCCGGCTGGCTGTTTATATGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 216 ArgIleLeuCysGlySerArgIlePheLeuProLeuThrArgLeuThrValIleLeuLeu 235
 Db 601 AGGATTCCTGTGTGATCCCGGAAGATGCGCTGACCGATGACCATCTCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluThrAlaLeuPheSer 255
 Db 661 ACAGTGTGCT 720
 QY 256 ArgIleHisLeuAspTrpIleValLeuPheCysHisValIleValSerIlePheLeu 275
 Db 721 AGGATCCACCTGATGTAAGTCTATTTGTGATGATGATGATGATGATGATGATGATG 780
 QY 276 SerAlaLeuAspSerSerAlaAspProIleIleThrPhePheValGlySerPheArgGlu 295
 Db 781 TCCGCTCTTAACAGACGAGCCCAACCCCATCTTCTCTCTGAGCTCTCTTAAAGCAG 840
 QY 296 ArgGluAsnArgGluAsnLeuIleValLeuValIleGluAlaArgAlaLeuGluIleAspThrProGlu 315
 Db 841 CGTCAAAATAGGCAAACTGAAAGCTGCTTCTCCAAAGGCTCTGACGACGACGCTGAG 900
 QY 316 ValAspGluGlyGlyIleThrLeuProGluIleThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGGATGAAGGTGAGGAGGTGCTCTCTCAAGAAACCTGAGAGTCTGGGAGACAGATTG 960
 QY 336 GluGlu 337
 Db 961 GAGCAG 966

RESULT 12
 ADO44605
 ID ADO44605 standard; DNA; 966 BP.
 AC ADO44605;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human HIT213 protein encoding DNA.
 XX
 KM HIT213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
 KM cytoskeletal; nephrotoxic; antiinflammatory; dermatological; analgesic;
 KM vulnerability; neuroprotective; human; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..966
 FT /*tag= a
 FT /product= "HIT213"
 XX
 PN WO2004039972-A1.
 PD 13-MAY-2004.
 PF 28-OCT-2003; 2003WO-JP013781.
 PR 29-OCT-2002; 2002JP-00314141.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA
 PI Katsuo Y, Watanabe T, Yasuhara Y, Mori I, Takeomi S;

XX
 DR WPI: 2004-376191/35.
 DR P-PSDB; ADO44604.
 XX
 PT HIT213 protein, encoded DNA and transgenic animals for clarifying
 PT pathological mechanism, developing therapeutic methods and screening
 PT preventives or remedies for related diseases e.g. cataract, cancer, and
 PT dermatitis.
 XX
 PS disclosure; SEQ ID NO 4; 161bp; Japanese.
 XX
 CC The invention relates to a non-human mammal that carries a DNA integrated
 CC with a foreign HIT213 or its mutant gene, or a part of it. The non-human
 CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
 CC onset, transient skin rash and proliferation-promoting activity. The
 CC foreign HIT213 gene is a gene that encodes a G protein-coupled receptor
 CC (GPCR) protein HIT213. The protein, its encoded DNA and constructed
 CC transgenic animals are useful for clarifying pathological mechanism,
 CC developing therapeutic methods and screening preventives or remedies for
 CC related diseases e.g. cataract, cancer, and dermatitis. The present
 CC sequence represents a DNA encoding a human HIT213 protein.

Alignment Scores:
 Seq Sequence 966 BP; 171 A; 280 C; 248 G; 267 T; 0 U; 0 Other;
 Pred. No.: 2,45e-148 Length: 966
 Score: 1683.00 Matches: 321
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.69% Mismatches: 0
 Query Match: 95.46% Indels: 0
 DB: 12 Gaps: 0

US-09-867-570-2 (1-337) x ADO44605 (1-966)

QY 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
 Db 1 ATGGATTCAACCATCCCACTTGGGTACAGAACTGACACCAATCAACGAGCGTAGGAG 60
 QY 36 ThrProCysTrpIleGluThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTGTCTACAGACGACCTGAGCTTACGAGGAGCTGACGAGATGATGTTCCCTGTGC 120
 QY 56 AlaLeuThrGlyAsnAlaValValLeuThrLeuLeuGlyCysArgMetArgArgAsnAla 75
 Db 121 GCGCTACAGAAACGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 76 ValSerIleTrpIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTACATCTCAACCTGGCGGCGGCACTTCTCTCTTACGCGCCACAT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 AATGTGTCGCGGTACGCTCAATCAATATCCCAATCCATCTCAAAATCTCACTCTCT 300
 QY 116 ValMetThrPheProIleArgIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATACCTTCTCCCACTTATAGGCTTAACACATGTCGAGCCATCAGACGAGGCGC 360
 QY 136 CysLeuSerIleLeuTrpProIleTrpIleGlyCysArgArgProArgTrpLeuSerSer 155
 Db 361 TGGCTGTCCATCTCTGGGCGCACTGTGTACCTGCGCGCCCAAGATACCTGTGCATCG 420
 QY 156 ValMetCysValLeuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluThrPhePhe 175
 Db 421 GTCATGTGTCTCTGCTGAGCCCTGCTCCCTGCTGCGAGATCTCTGAGTGAAGTTC 480
 QY 176 CysAspPheLeuPheSerGlyValIAspSerValITPcYsgluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTCTGTAGTGTCTGATTTCTGTTGGTGTGAACGTCGATTTTCATTACA 540
 QY 196 IleAlaTPLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCCGGCTGGCTGTTTATATGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY	276	SeqlaLeuAnSeSerSlaaAnPoliEiEYrPhaPheValGlySerPheArgGln	295
Db	781	TCGGCTCTTTACAGCAGATGCCAACCCATCATTTACTTCTTCGTGGGCTCTTTAGGCG	840
QY	296	ArgGlnaAnArgGlnaAnLeuLysIeuValLeuGlnArgAlaLeuGlnaAnPThrProGln	315
Db	841	CGTCAAAATVAGGCAGACCTGGAAGCTGTTCTTCACAGGGCTCTGCAGGACACGCTTGAG	900
QY	316	ValaAnPGLuGLyGLyGLyYTPLeuProGlnGlnIuThLeuGlnLuleuSerGlySerArgLeu	335
Db	901	GTGATGTGAAGGTGAGGGGTCTTCTCAGAGAAACCTGAGCTGTGGGAGACAGATTG	960
QY	336	Glugin	337
Db	961	GAGCAG	966
RESULT 14			
ADP29104			
ID	ADP29104	standard; cDNA; 969 BP.	
XX	ADP29104;		
AC			
XX			
XX	12-FEB-2004	(first entry)	
DE		Human GPCR R-14-encoding cDNA.	
XX			
KW	Human, R-14; GPCR; G protein coupled receptor;		
KW	iridectomy meshwork tissue; ocular outflow; antagonist;		
KW	ocular hypotensive; drug screening; elevated intraocular pressure;		
KW	glaucoma; ophthalmological; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003080659-A1.		
XX			
PD	02-OCT-2003.		
XX			
PF	27-MAR-2003; 2003WO-CA000444.		
XX			
PR	27-MAR-2002; 2002US-0367513P.		
XX			
PA	(THER-) THERATECHNOLOGIES INC.		
XX			
PI	Peri KG, Moffett S, Abran D;		
DR	WPI; 2004-053019/05.		
XX	P-PSDB; ADP29105.		
PT	New substantially pure R-14 polypeptide, useful as drug targets for		
PT	lowering intraocular pressure and for treating condition such as		
PT	glaucoma.		
PS	Claim 22; SEQ ID NO 1; 85bp; English.		
XX			
CC	The invention relates to a human GPCR (G protein coupled receptor)		
CC	designated R-14 (ADP29105), nucleic acids encoding it (ADP29104), and R-		
CC	14 peptide antagonists (ADP29106-ADP29108). The invention also		
CC	encompasses vectors and host cells comprising R-14 nucleic acids, and		
CC	methods of screening for R-14 antagonists. The R-14 receptor is expressed		
CC	in trabecular meshwork tissue, and is associated with a role in ocular		
CC	outflow. Inhibition of the receptor results in a reduction of basal		
CC	intraocular pressure, making the R-14 receptor a useful target for		
CC	screening for ocular hypotensive drugs. The R-14 peptide antagonists are		
CC	useful for reducing intraocular pressure for the treatment of conditions		
CC	associated with elevated intraocular pressure such as glaucoma and		
CC	related conditions. The R-14 receptor, and host cells expressing an R-14		
CC	polynucleotide, may be used in screening for R-14 receptor antagonists.		
CC	The present sequence is related to the invention.		
XX			
SQ	Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;		
Alignment Scores:			

Pred. No.:	2,466-148	Length:	969
Score:	1683.00	Matches:	321
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	95.46%	Indels:	0
DB:	12	Gaps:	0

US-09-867-570-2 (1-337) x ADP29104 (1-969)	
QY	16 MetApSerThrIleProValLeuGIyThrGIuLeuThrProIleAsnGIyAArgGIuGIu 35
Db	1 ATGATGATCAACCATCCCACTCTTGGGTACAGAACTGACCAACCAACGACCGTAGAG 60
QY	36 ThrProCySerTrIyGlnIhrIleuSerPheThrGIyLeuThrCysIleAlSerLeuVal 55
Db	61 ACTCTTGGTACAGAGACCCCTGAGCTTCACGGGGCTGAGGTTCATCTTCCCTTGTG 120
QY	56 AlAluThrGIyAsnAlaValIleuThrIleuLeuGIyCysArgMetArgAsnAla 75
Db	121 GCCGTGACAGAGAACCGGCTGTGTGCTCGGCTCCGGGCTGCGCATGCCGAGGACGCT 180
QY	76 ValSerIleTrIleuAsnLeuValAlaAlaPheLeuPheLeuSerGIyHisIle 95
Db	181 GTCTTCATCTACATCTCTCAACCTGGTCGGGCGGCACTTCCCTTCCCTTAACGGGCAATT 240
QY	96 IlCySerProLeuArgIleuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db	241 ATATGTTGCCGTTAGCGCTCATCAATTCGGCATCCCATCTCCAAATTCCTCAGCTC 300
QY	116 ValMetThrPheProTrpPheIleGIyLeuSerMetLeuSerAlaIleSerThrGIuArg 135
Db	301 GTGATGACCTTTCCTACTTATTAAGCGCTTAAGCATGCTGAGCGCATCAGACCGAGCGC 360
QY	136 CysLeuSerIleLeuTrpProIleTrpTrpIleCysArgAsnArgProArgTrpLeuSerSer 155
Db	361 TGCCTGTCACCTCTGAGGCCATCTGGTACACATGCGCGCGCCGACATTAACCTGTCAATCG 420
QY	156 ValMetCysValIleuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGIuTrpMetPhe 175
Db	421 GTCATGTGTCTCTGCTCTGGGCGCTGTCCCTGCTCGCGGATATCTCGAGTGGATGTTTC 480
QY	176 CysAspPheLeuPheSerGIyAlaAspSerValTrpCysGIuThrSerAspPheIleThr 195
Db	481 TGTGACTTCTGTTTGTAGTGCTGCTGATCTCTGTTGTGTGAAGCGTCAGATTTCATTACA 540
QY	196 IleAlaTrpIleuValPheLeuCysValIleuLeuCysGIySerSerSerLeuValIleuVal 215
Db	541 ATCGCGGTGGTGTATTTTATGTGTGTGTCTCTGTGTGGGTCCACCGCTGTCTGTGTGTC 600
QY	216 ArgIleLeuCysGIySerArgIySmetCProLeuThraGleuTrpValThrIleLeuLeu 235
Db	601 AGATTCTCTGTGTGATCCCGGAAGATGCCGCTGACACAGCGGTGACGATCATCTCTC 660
QY	236 ThrValIleuValPheLeuLeuCysGIyLeuProPheGIyIleGIuTrpAlaLeuPheSer 255
Db	661 ACAGTGTGTCTCTCTCTCTGTGGCGCCCTTGTGATTCAGTGGGCCCTGTTTTCC 720
QY	256 ArgIleHisLeuAspTrpIlyValIleuPheCysHisValHisLeuValSerIlePheLeu 275
Db	721 AGGATTCACCTGGATTTGAAAGCTTATTTGTCTACATGTGACATCTAGATTTCTCTG 780
QY	276 SerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGIySerPheArgGln 295
Db	781 TCCGCTCTTAACGACGATGCCAACCCCATTTACTTTCGTGGGGCTCTTTAGGCAG 840
QY	296 ArgGlnAsnArgGlnAsnLeuIlyLeuValIleuGIuArgAlaLeuGlnAspTrpProGln 315
Db	841 CGTCAAAATATGGGACAACTCGAAGCTGTCTTCCAGAGGGCTCTGACAGGACACCCCTGAG 900
QY	316 ValAspGIuGIyGIyTrpIleuProGIuGIuThrIleuGIuLeuSerGIySerArgLeu 335
Db	901 GTGATATAGGTGAGGGGTCTTCTCAGAGAAACCTTGAGACTGTCTGGGAGACAGATTG 960

QY	336	Gluc1n	337
Db	961	GAGCAG	966
	/		
	RESULT 15		
ID	AAA70342	standard; cDNA; 969 BP.	
XX	AAA70342		
AC	AAA70342;		
XX			
DT	19-DEC-2000	(first entry)	
XX			
DE	Human novel G-protein coupled receptor #1 coding sequence.		
XX			
KM	Human; novel G-protein coupled receptor; signal transduction;		
KW	disease diagnosis; drug screening; disease therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CD5	1..969	
FT		/*lag= a	
FT		/product= "novel G-protein coupled receptor #1"	
PN	W0200040724-A1.		
XX			
PD	13-JUL-2000.		
XX			
PF	04-JAN-2000; 2000MO-US000052.		
XX			
PR	04-JAN-1999; 99US-0114666P.		
PR	14-JAN-1999; 99US-0115828P.		
XX			
PA	(LEXI-) LEXICON GENETICS INC.		
XX			
PI	Nehls M, Wadtler F;		
XX			
DR	WPI; 2000-465986/40.		
DR	P-PSDB; AAB14846.		
PT	New polynucleotides encoding novel G-protein coupled receptors useful for		
PT	diagnosis, drug screening, clinical trial monitoring and for the		
PT	treatment of physiological or behavioral disorders.		
XX			
PS	Claim 1; Page 52-53; 61pp; English.		
CC	The present sequence is the coding sequence for a novel human G-protein		
CC	coupled receptor (NGPCR). These proteins are involved in signal		
CC	transduction pathways in many cases. The protein contains seven		
CC	transmembrane domains, and is expressed in human testis, mammary gland		
CC	and salivary gland tissues. The gene, its protein, agonists, antagonists		
CC	and antibodies can be used to diagnose and treat diseases associated with		
CC	the inappropriate expression or expression of mutant versions of the		
CC	protein, for screening for drugs which can be used in the same manner,		
CC	and for elucidating the function of the protein		
XX			
SQ	Sequence 969 BP; 172 A; 282 C; 249 G; 266 T; 0 U; 0 Other;		
	Alignment Scores:		
	Pred. No.:	1,86e-146	Length: 969
	Score:	1663.00	Matches: 320
	Percent Similarity:	99.38%	Conservative: 0
	Best Local Similarity:	99.38%	Mismatches: 2
	Query Match:	94.33%	Indels: 0
	DB:	3	Gaps: 0
	US-09-867-570-2 (1-337) x AAA70342 (1-969)		
QY	16	MeAspSerThrIleProValleuGlyThrGluLeuThrProIleAsnGlyArgGluGlu	35
Db	1	ATGATTCAACCACTCCACGCTTGCGGTACGAGCACTACCAATCAACGACGTGAGAG	60
QY	36	ThrProCysTyrIlysglnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal	55

Db	61	ACTCTTGCTACAGACAGACCCCTAGCTTCAAGGGGCTGACGTGATCGTTTCCCTGTTC	120
Qy	56	AlaLeuThrGlyAsnAlaValValLeuTrpLeuGlyCysArgMetArgAsnAla	75
Db	121	GGCGTGAAGAGAAACGGCGTTGTGCTCTGGGCTCCCTGGGCTGCGGATGCGCAGAAACGCT	180
Qy	76	ValSerIleTyrIleLeuAsnLeuValAlaAspPheLeuPheLeuSerGlyHisIle	95
Db	181	GTCCTTCATCTACATCTCAACCTGGTCGGGGCCGACCTTCTTCTTTAGCGGCCACATT	240
Qy	96	IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIlyValLeuSerPro	115
Db	241	ATTAGGTTGCGCGTTACGGCTCATCAATATCCGACATCCCATCTCCAAATCTCAATCTCT	300
Qy	116	ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg	135
Db	301	GTGATGACCTTTCCTACTTATAGCCCTTAGCAATGAGGCATGACGCATGACACCGAGCC	360
Qy	136	CysLeuSerIleLeuTrpProIleTyrPyrHisCysArgArgProArgTyrLeuSerSer	155
Db	361	TGCCTGTTCATCTCTGTGGCCCATCTGGTACACATCGCCGCCGACAGTACCTGTATCG	420
Qy	156	ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleGluIleTrpMetPhe	175
Db	421	GTCAATGTGTGCTGTGCTCTGGGCGCTGTGCCCTGCTGCGAGATACCTGAGAGTGAATGTC	480
Qy	176	CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThr	195
Db	481	TGTACCTTCCGTGTATGAGTGCGTCAATCTGTTGGTGTGAACGTCACATTCATTAC	540
Qy	196	IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal	215
Db	541	ATCCGGTGGCTGTTTTTTATGTGTGTTCTGTGGGTCACGCTGGTCTCGTGGTTC	600
Qy	216	ArgIleLeuCysGlySerArgIysMetProLeuThrArgLeuTyrValThrIleLeuLeu	235
Db	601	AGGATTCCTCTGTGTGATCCCGGAAGATGCCGTGACAGGCTGTAGTCATCTCCCTTC	660
Qy	236	ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSer	255
Db	661	ACAGTGTGGTCTTCTCTCTCTGTGGCTTGCCCTTTGGCATTCAGTGGGCTCTGTTTTC	720
Qy	256	ArgIleHisLeuAspTrpIysValLeuPheCysHisValHisLeuValSerIlePheLeu	275
Db	721	AGGATCCACCTGGATTTGGAAGTCTTATTTGTGCATGTGCATCTAGTTCCATTTTCC	780
Qy	276	SerAlaLeuAsnSerSerHisAsnProIleIleTyrPhePheValGlySerPheArgGlu	295
Db	781	TCCGCTCTTAAACAGACAGTGCACACCCCAATCACTTCTTGTTGGGCTCTTTTAGGACG	840
Qy	296	ArgGluAsnArgIleAsnLeuIysLeuValLeuGluArgAlaLeuGluAsnThrProGlu	315
Db	841	CGTCAAAATAGGCAGAACTGAAGCTGTTCTCCAGAGGGCTCTGCAGGACACGCTTAG	900
Qy	316	ValAspGluGlyGlyTyrTrpLeuProGluGluThrLeuGluLeuSerGlySerArgLeu	335
Db	901	GTGATGTAAGGTGAGGGGCGACCTTCTCAGGAAACCTCTGAGAGCTGCGGGAAGCAGATTG	960
Qy	336	GluGlu 337	
Db	961	GAGCAG 966	

Search completed: December 1, 2004, 23:05:59
Job time : 505 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 1, 2004, 22:57:34 ; Search time 2974 Seconds
(without alignments)
4129.183 Million cell updates/sec

Title: US-09-867-570-2
Perfect score: 1763
Sequence: 1 MESKSWVIRLGLSLMDSTI.....EGGGWLPQRTLEHSGNRLEQ 337

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xip
-Q=/cgm1/USPTO.spool.p/US09867570/runat.01122004.161123.6350/app.query.fasta_1.519
-DB=EST -OPMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09867570 @CGN 1.1 5180 @runat.01122004.161123.6350 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1221	69.3	889	6	CA455045 AGENCOURT
2	1039.5	59.0	889	7	CN835944 AGENCOURT
3	1032.5	58.6	764	4	BG198766 RST18035
4	1003.5	56.9	688	2	BR439409 HTM1-025F
5	978	55.5	825	7	CN843258 AGENCOURT
6	976	55.4	835	7	CN843633 AGENCOURT
7	971.5	55.1	827	5	CN831885 AGENCOURT
8	966	54.8	924	7	CN839634 AGENCOURT
9	898.5	51.0	787	4	BG208126 RST7616

10	853	48.4	871	7	CN839125 AGENCOURT
11	820	46.5	924	6	CD051096 AGENCOURT
12	815	46.2	871	4	BG210740 RST30287
13	772	43.8	953	7	CN839765 AGENCOURT
14	770	43.7	1009	7	CN837763 AGENCOURT
15	763	43.3	2075	3	AK029369 Mus muscu
16	612	34.7	503	8	AO888076 HS 2142 B
17	610	34.6	468	8	B74348 CTT-HSP-204
18	591	33.5	699	9	CR190735 Forward s
19	588.5	33.4	620	8	AO696198 HS 2142 B
20	567.5	32.2	905	6	CD243595 AGENCOURT
21	558.5	31.7	811	7	CN831931 AGENCOURT
22	550.5	31.2	704	6	BY741352 BY741352
23	542	30.7	895	7	CN840024 AGENCOURT
24	539	30.6	682	8	AZ346449 AGENCOURT
25	533.5	30.3	788	9	CR169980 Reverse s
26	530	30.1	733	6	BY741303 BY741303
27	505	28.6	3575	3	AK048528 Mus muscu
28	498.5	28.3	777	6	CD467625 Leuko2_1
29	498.5	28.3	783	6	CD469424 Leuko2_1
30	494.5	28.0	770	6	CD469131 Leuko2_1
31	493.5	28.0	618	4	B1828553 603078443
32	490.5	27.8	761	6	CD467287 Leuko2_1
33	483.5	27.4	765	6	CD469412 Leuko2_1
34	480.5	27.3	756	6	CD469433 Leuko2_1
35	478	27.1	725	6	CD468987 Leuko2_1
36	477	27.1	515	1	AL705589 DKE2p686K
37	471	26.7	503	1	AF003828 AF003828
38	469.5	26.6	722	7	CO565764 AGENCOURT
39	466	26.4	954	7	CN842891 AGENCOURT
40	461	26.1	484	2	BB853701 BB853701
41	456.5	25.9	576	6	CD469125 Leuko2_1
42	454	25.8	719	6	CD469292 Leuko2_1
43	452.5	25.7	2393	3	AK030261 Mus muscu
44	452.5	25.7	2404	3	AK030265 Mus muscu
45	451	25.6	704	6	BY741329 BY741329

ALIGNMENTS

RESULT 1
CA455045
LOCUS
DEFINITION
CA455045
AGENCOURT_10735784 MAPL Homo sapiens CDNA clone IMAGE:6722551 5',
mRNA sequence.
ACCESSION
CA455045
VERSION
CA455045.1 GI:24905369
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
Nih-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 889)
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLM4285 row: h column: 07
High quality sequence stop: 743.
Location/Qualifiers
1. 889

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722551"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HMEL, LNCaP"
/lab_host="EMD10B"
/clone_lib="MAPCL"
/note="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subcloned with brain, liver, lung, kidney and muscle.
directionally cloned. Priming method: oligo-dt. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kirsti A. Eglund, James J. Vincent, Robert Strauberg,
Bungkok Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Alignment Scores:
Pred. No.: 2,566-114 Length: 889
Score: 1221.00 Matches: 237
Percent Similarity: 98.34% Conservative: 0
Best Local Similarity: 98.34% Mismatches: 3
Query Match: 69.26% Indels: 1
DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x CA455045 (1-889)

QY 8 Val11eAgtleuGlyPheleuSerMetApsSerThr11eProVal1euglyThrg1uleu 27
Db 149 GTGATCAGACTGGGGTTTCTGACATGATTCACACCTCCAGTCTGGGTACAGAACTG 208
QY 28 ThrPro11easnglyAgt1eug1uthrProCyTy11eag1inThrl1euserpethr1g1y 47
Db 209 ACACCAATCAACGAGAGTACGAGAGTCTGCTACAGACAGACCTGAGCTTCAAGGAG 268
QY 48 LeuThrCy11e11eValSer1e1uVal11e1eThrg1yAAsn11eVal11e1eThr11e 67
Db 269 CTGAGCGATGATGTTCTCTTCTGCGCTGACGAGAAAGCGGTTGCTGCTGCTGCTG 328
QY 68 GlyCyAgtMetAgtAgtAAsn11eValSer11e11eThrl1e1e1e1e1e1e1e1e1e 87
Db 329 GGGTCCCGCATCGCAGAAACGCTGCTCTCATGCTGATCTCACTCACTGCTGCTGCTG 388
QY 88 Phe1e1uPhe1e 107
Db 389 TTCCCTCTCTTACGCGGACCATATATGTTCCCGTACGCTCACTCACTCACTCACTCA 448
QY 108 Pro11eSer11e11e 127
Db 449 CCCATCTCCAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 508
QY 128 LeuSer11e11e 147
Db 509 CTGACCGCATGACGACCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 568
QY 148 ArgAgtProAgtTy11e 167
Db 569 CGCCCCCAGATACCTGTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
QY 168 ArgSer11e 187
Db 629 CGGATATCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
QY 188 CysG1uThrSerApsPhe11eThrl1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e 207
Db 689 TGTGAAGCTCAGATTTTCAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
QY 208 GlySerSer1e 227
Db 749 GGGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
QY 228 Arg1e1u11e 247
Db 809 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868
QY 247 e 247

Db 869 T 869

RESULT 2

CN835944/c 863 bp mRNA linear EST 02-JUN-2004
LOCUS AGENCOURT 15864324 NIH_MGC.145 Homo sapiens cDNA clone
DEFINITION IMAGE:30706618 3', mRNA sequence.

ACCESSION CN835944

VERSION CN835944.1 GI:47941599

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 863)

AUTHORS NIH-MGC http://mgs.mcl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contract: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics / NIH

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: c9apbs-r@mail.nih.gov

Tissue Procurement: GPCR Consortium

cDNA Library Preparation: GPCR Consortium

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNLN at:

http://image.llnl.gov

Plate: IRB14 row: e column: 11

High quality sequence stop: 719.

Location/Qualifiers

1. 863

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30706618"

/issue_type="mixed"

/lab_host="DH10B"

/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:

varies by clone; ORFs were PCR-amplified and cloned into

pCDNA3.1 by the GPCR Consortium. Cloning sites vary by

clone and include the following: 5'-EcoRV-XbaI/XhoI-3',

5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional).

For information about which gene each clones represents,

please visit our anonymous ftp site at

ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat

A Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 9,856-96 Length: 863

Score: 1039.50 Matches: 211

Percent Similarity: 86.42% Conservative: 18

Best Local Similarity: 79.62% Mismatches: 33

Query Match: 58.96% Indels: 3

DB: 7 Gaps: 1

US-09-867-570-2 (1-337) x CN835944 (1-863)

QY 74 AAsn11eValSer11eTy11e11e1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e 92
Db 863 AACGCTGCTCCATCTACATCTTCAACCTGCGCCGACGAGACTTCTTCTCTTCA 804
QY 92 eGlyHis11e11e 112
Db 803 GCTTCCAGATTTACCTTGCATTTACGCTGATGATGATGATGATGATGATGATGATG 744
QY 112 1e 132

[illegible]

Email: scatinateherys.com
 High quality sequence stop: 559.
 Location/Qualifiers
 source
 1..764
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_id="Atherys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,276-95 Length: 764
 Score: 1032.50 Matches: 205
 Percent Similarity: 86.77% Conservative: 18
 Best Local Similarity: 79.77% Mismatches: 29
 Query Match: 58.56% Indels: 5
 Db: 4 Gaps: 1
 US-09-867-570-2 (1-337) x BG198766 (1-764)
 Oy 78 ILeTyrIleLeuAsnLeuValAlaAlaAspPheLeuSer-GlyHisIleIleCy 97
 Db 764 ATCTACATCTCTACATGATGCGCGGACAGACTTCTCTGCTCAGCGGCGCCTTATATA 705
 Oy 97 sSerProLeuArgLeuIleAsnIle-ArgHisProIleSerIlyIleLeuSerProValM 117
 Db 704 TTCCCTGTACTCTTCATCAGATGATCCCCCATACATCTTAATAATCCCTTATCCGTGA 645
 Oy 117 eCtIhrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysL 137
 Db 644 TGATGTTTTCATCTTCAGCGCCGTGAGCTTTCTAGTGCCGTGAGCAGCAGCGCTGCC 585
 Oy 137 euserIleLeuTrpProIleTrrPyrHisCysArgArProArgTyrLeuSerSerValM 157
 Db 584 TGTGCGTCTGTGGCCATCTGTATACCGGTGCCACGCCCCACACACACTGTACCGGTGG 525
 Oy 157 eCysValleuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysA 177
 Db 524 TGTGTGCTCTCTCTGGGCCCTCTGCTCCCTGCGAGACATCTGGAGTGTATATGTG 465
 Oy 177 sPheLeuPheSerGlyAlaAspSerValTrrPyrGlyIhrSerAspPheIleThrIleA 197
 Db 464 GCTTCCTGTACAGTGCTGTATTTGCTGTGTGTCAACATCAGATTTCATCAGCTGCG 405
 Oy 197 lAtPheLeuValPheLeuCysValIleLeuCysGlyIserSerLeuValLeuLeuValArgI 217
 Db 404 COTGGCTGATTTTATATGTGTGTCTCTGTGTGGGCCAGCCGGCTCGTGTACAGA 345
 Oy 217 lLeuCysGlyIserArgIysMetPLeuLeuThrArgLeuTyrAlaTrrIleLeuLeuThyV 237
 Db 344 TTCCTGTGTGATCCCGGAAGATACCGCTGACACAGCTGTACGTGACATCTGCTACAG 285
 Oy 237 alLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrrAlaLeuPheSerArgI 257
 Db 284 TACTGTCTTCTCTCTCTGTGGCCGTGCCCTTTGGCATTCAGTTTTCCTATTTTATNGA 225
 Oy 257 lHisIleAsnArTrrPlyIysValLeuPheCysHisValHisLeuValSerIlePheLeuSerA 277
 Db 224 TCCAGGTGACAGGGAAGCTTATTTTGTGCAGTTCATCTAATTTCTATTTTCTGTCCG 165
 Oy 277 lAlLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGluArgG 297
 Db 164 CTCCTTAACAGACGTCCCAACCCCATCTTATCTTCTGTGGGCTCTTTAGCAGCGTC 105
 Oy 297 lAsnArgGlnAsnLeuIysLeuValLeuGlnArgAlaLeuGlnAsnThrProGluValA 317

Db 104 AAAATAGGACAGAACCTGAGTCTTCACAGAGGGCTCTGACAGACGGCTGAGGTG 45

Qy 317 spglluglvglyglttrpleuProGlnGluThrlleuGluSerGly 332

Db 44 ATGAAGGTGAGGAGCAATCTCTGAG-----GAAATGCTTGA 7

RESULT 4

BE439409 688 bp mRNA linear EST 25-JUL-2000

LOCUS HTML-025F1 HTML Homo sapiens cDNA, mRNA sequence.

DEFINITION BE439409

ACCESSION BE439409.1 GI:9438891

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Gonzalez, P., Epstein, D.L. and Borras, T.

TITLE Characterization of gene expression in human trabecular meshwork using single-pass sequencing of 1060 clones

JOURNAL Invest. Ophthalmol. Vis. Sci. (2000) In press

COMMENT Contact: Pedro Gonzalez

Department of Ophthalmology

Duke University

Duke Eye Center, Erwin Rd, Box 3802, Durham, NC 27710-3802, USA

Tel: 919 681 4085

Fax: 919 684 8983

Email: pedro.gonzalez@duke.edu.

Location/Qualifiers

1..688

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="trabecular meshwork"

/clone_lib="HTML"

ORIGIN

Alignment Scores:

Pred. No.: 3,42e-92 Length: 688

Score: 1003.50 Matches: 211

Percent Similarity: 96.80% Conservative: 1

Best Local Similarity: 96.35% Mismatches: 7

Query Match: 56.92% Indels: 6

DB: 2 Gaps: 0

US-09-867-570-2 (1-337) x BE439409 (1-688)

Qy 8 ValileargleuGlyPheleuserMetacpserThrlleProValleuGlyThrlleu 27

Db 36 GTCAACAGACTGGGGTTCTGAGCATGATCAACATCCCATCTTGGGTACAGAACTG 95

Qy 28 ThrProillearglylYargGluGluThrProCyTYrlysglInThrlleuserPhehrgly 47

Db 96 AACACCAATCAACGAGCGTGAAGAGACTCTTGCTACAGACAGACCTGAGCTTCAAGGG 155

Qy 48 leuThrCyrlleValSerleuValAlaleuThrlYAsnalAlaValleuThrlleu 67

Db 156 CTGAGGTGATGTTCTTCCTTGGCGGTGACGAGAAACGGGTGTGCTGCGCTCTG 215

Qy 68 GlyCyArGMeArGArGAsnaAlaValSerlleYrleleuAnleuValAlaAlaAsp 87

Db 216 GGCTGCCGATCCGACGAAACCTGTCTCATTTACATCTCAACTGTCGGGGCGGAC 275

Qy 88 PheleuPheleuserGlyhlslelleCySerProleuArgleuileanllearghls 107

Db 276 TTCCTCTTCCTTACGGGCGACATTAATATGTTCCCGCTTACGCTCAATATCCGCAT 335

Qy 108 ProileSerlelleleuserProValMetThrlPheProYrPheilleGlyleuserMet 127

Db 336 CCCATCTCCAAATCTCATGCTCTGATGACCTTTCCTACTTATAGGCCCTTAAGCATG 395

Qy 128 leuserAlalleSerThrlleuArgCyLeuserlleleuThrlProlleThrlleCyS 147

Db 396 CTGAGGGCATCAGACCGAGGCTGCCCTGTCATCTGTGGCCATCTGGATCAGCTGC 455

Qy 148 ArgArgProArgTYrleuserSerValMetCyValleuThrlleuPalaSerleu 167

Db 456 CGCCGCCAGATACCTGTGCATGCGTGCATGTGTCTGTCTGTGGCCCTGTCTGTG 515

Qy 168 ArgSerlleleuGluThrlPheCyAspPheleuPheSerGlyAlaAspSerValTrp 187

Db 516 CGAGATCTCTGAGATGATGTTCTGTGACTTCTTGTAGTGAGCTAATTCGTTGG 575

Qy 188 CyGluThrSerAspPheilleThrlleAlaTrpleuValPheleuCyValleuCyS 207

Db 576 TGTGA-ACGTCAATTTATTTAATCAATCGGTGCTGAT-TTTTATGTGTGT-CTCTGT 632

Qy 208 GlySerSerleuValleuValArglleuCyGlySerArglyMetProleu 226

Db 633 GGCTCCAG-CTGTCTCTGTGTGTGAGAT-CTCTGTGATC-CGAGATGACAGCTG 686

RESULT 5

CN843258/c 825 bp mRNA linear EST 02-JUN-2004

LOCUS AGENCOURT.15669738 NIH_MGC_145 Homo sapiens cDNA clone

DEFINITION CN843258

ACCESSION CN843258

VERSION CN843258.1 GI:47948913

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10N07 Bethesda, MD 20892

Email: cga@bbs-remail.nih.gov

Tissue Procurement: GPCR Consortium

cDNA Library Preparation: GPCR Consortium

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Plate: IRB14 row: e column: 03

High quality sequence start: 21

High quality sequence stop: 561.

Location/Qualifiers

1..825

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30706610"

/tissue_type="mixed"

/lab_host="DH10B"

/clone_lib="NIH_MGC_145"

/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/NotI-3', EcoRV (3A cloned, non-directional). For information about which gene each clone represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/tearayed_plates/IRB1.presv.dat

A Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 1.85e-89 Length: 825

Score: 978.00 Matches: 203

Percent Similarity: 86.33% Conservative: 18
 Best Local Similarity: 79.30% Mismatches: 29
 Query Match: 55.47% Indels: 7
 Gaps: 3

US-09-867-570-2 (1-337) x CN843258 (1-825)

QY 85 AAAAAAAAAA...leuPheleuSerGlyHisIleCysSerPro...leuAtgLeuI 103
 DB 794 GCGACAACTTCCTCTCTCCACGGGCGCTTATTATTATTCCTCCGGTTAAGGTTCCA 735
 QY 103 eaaNilArg---HisProIleSerLys---IleLeuSerProValMetThrPheProY 121
 DB 734 TNCAGTATCCCGCCCACTCCATCTTAAATCTTCATC-CCTGTGATGATGTTTCTAC 676
 QY 121 rpehIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuT 141
 DB 675 CCGGACGAGCTGAGCTTTTGAAGTCCGTGACCGAGCGCTGCTGCTCCCTG 616
 QY 141 pProIleTrpTyHisCysArg-ArgProArgTyIleuSerSerValMetCysValleu 161
 DB 615 GCCCATCTGTGATCCGCTGCCACACGCCCACTGTCAAGCGGTGTGTCTCTGC 556
 QY 161 eutTPAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeu 181
 DB 555 TCTGGAGCTGTCTCTCTGCGAGCATCTCGAATGATGATGATGCTCTCTCA 496
 QY 181 erGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValP 201
 DB 495 GTGGTGTGATTCCTGCTGGTGTCAAAATCAGATTCATCAGTCGGCTGCTGATTT 436
 QY 201 heLeuCysValValleuCysGlySerSerLeuValIleLeuValArgIleLeuCysGly 221
 DB 435 TTTTATGTGTGCTCTGTGGGTCCAGCTGTCTGTGATCAGATTCCTGTGAT 376
 QY 221 eArGlyMetProLeuThrArgLeuTyValThrIleLeuLeuThrValleuValPhe 241
 DB 375 CCGGAAAGATACCGCTACGCTGATGACATCTGTCTACAGTATGCTCTCC 316
 QY 241 euleuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 261
 DB 315 TCCCTGTGGCGCTGCTGGTGGCATTCAGTTTTCATTTTATGATCCAGCTGACA 256
 QY 261 rplysValleuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 281
 DB 255 GGGAAAGCTTATTTGTCATGTCATGATTCATTTCTATTTCTGCTTAAACAGCA 196
 QY 281 eAlaAsnProIleIleTyPhePheValGlySerPheArgGlnArgGlnAsnArgGln 301
 DB 195 GTGCGAACCCCATCATTTACTTCTTCGCGCTCTTGAAGCAGCGTCAAAATAGGACA 136
 QY 301 anLeuValleuValleuGlnArgAlaLeuGlnAspThrPro-GluValAspGluGly 320
 DB 135 ACCGGAAGCTGCTTCTCAGAGGCTCTGACAGAGCGGTCCAGAGTGAAGAGTGA 76
 QY 321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArg 334
 DB 75 GGGCAGCTCTCTGAGAAATCTGAGAGCTGTGGGAAAGACA 34

RESULT 6
 CN843633/c 835 bp mRNA linear EST 02-JUN-2004
 LOCUS AGENCOURT_15658848 NIH_MGC_145 Homo sapiens cDNA clone
 DEFINITION IMAGE:30706619 5', mRNA sequence.
 ACCESSION CN843633
 VERSION CN843633.1 GI:47949288
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 835)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: egabbs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 CDNA Library Preparation: GPCR Consortium
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: IRB14 row: e column: 12
 High quality sequence stop: 660.
 Location/Qualifiers
 1. 835
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706619"
 /issue_type="mixed"
 /lab_host="DH10B"
 /clone_1lb="NIH_MGC_145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
 varies by clone; ORFs were PCR-amplified and cloned into
 pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
 clone and include the following: 5'-EcoRV-XmnI/XhoI-3',
 5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional).
 For information about which gene each clone represents,
 please visit our anonymous ftp site at
 ftp://image.llnl.gov/image/rearrayed_plates/IRB14.presv.dat
 a Note: this is a NIH_MGC Library."

FEATURES
 source

ORIGIN

Alignment Scores:
 Pred. No.: 3,036-89 Length: 835
 Score: 976.00 Matches: 198
 Percent Similarity: 86.85% Conservative: 20
 Best Local Similarity: 78.88% Mismatches: 30
 Query Match: 55.36% Indels: 6
 DB: 7 Gaps: 1

US-09-867-570-2 (1-337) x CN843633 (1-835)

QY 88 PheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHis 107
 DB 822 TTCCTTCAAGCTTTCCAA-GATTATACGTTTCCGCAATACGCTTTATCATATTT-CAG 765
 QY 108 ProIleSerLysIleLeuSerPro-----ValMetThrPheProTyIlePheIleGly 124
 DB 764 CCATCTCATCCGCGAAG-AATCCCTGCTTTCGTGATGACCTTCCCTTACAGGC 706
 QY 125 LeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrp 144
 DB 705 CTAGTATGTGTGAGCGGCATCAGACCGAGCGTCTGTCTGTCTGTGAGCCCATCTGG 646
 QY 145 TyrHisCysArgArgProArgTyIleuSerSerValMetCysValleuLeuTrpAlaLeu 164
 DB 645 TACCGTCCCGCGCGCCACACACCTGTCCAGCGGTCTGTGTCTCTGCGGCGCTG 586
 QY 165 SerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAsp 184
 DB 585 TCCCTGTGTTTGTATGATGAGTGAAGTTCGTGATCTTCCTGTTAAGGTCGAT 526
 QY 185 SerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysVal 204
 DB 525 TCTAGTGTGTGTAAAGCTCAGATTCATCCAGTCCCGGCTGATTTTATATGTGTG 466
 QY 205 ValLeuCysGlySerSerLeuValIleLeuValArgIleLeuCysGlySerArgValMet 224
 DB 465 GTTCTGTGTCTTCCAGCTGTGCTCTGTGTGATCCTCTGTGATCCCGAAGATG 406

QY	225	ProleutrraAglyLeuTyValThrIleLeuLeuTrValIleValPheLeuLeuCy61y	244
Db	405	CCGCGACCAAGCCTCTACGTACCACTCCTGCTCAAGTCGTCGTCTCTCTCTCGCGGC	346
QY	245	LeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIlyValIleu	264
Db	345	CTGCCCTTCGGCATTCCTGGGGGCCCTTAATTTCACAGATGCACCTTAATTGGAAAGCTCTTA	286
QY	265	PheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAspSerPheAlaAsp	284
Db	285	TATTCATGATTAATCTGATGCTTTGCAATGCCCTGCTCTCTTAACAGATGTCACACCC	226
QY	285	IleIleTyPhePheValAlGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuIlySeru	304
Db	225	ATCATTTACTCTTCTGTCGGGCTCCTTTAGCGACGCTCAAAATAGGCAACACTGAAGCTG	166
QY	305	ValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTrpLeu	324
Db	165	GTTCCTCCAGAGGCTCTGACAGACAAAGCTGAGTGATTAAGGTGAAGGACGCTTCT	106
QY	325	GlnGluThrLeuGlnLeuSerGlySerArgLeu	335
Db	105	GAGGAAGCTTGAGACTGTGGGAAAGACAGATTG	73
RESULT 7			
LOCUS	CN831885/c	827 bp	mRNA linear EST 02-JUN-2004
DEFINITION	AGENCOURT.15669850 NIH_MGC.145 Homo sapiens cDNA clone		
IMAGE:	30706617 5', mRNA sequence.		
ACCESSION	CN831885		
VERSION	CN831885.1	GI:47935638	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteleostomi; Primates; Catarrhini; Homidae; Homo.		
TITLE	1 (baee 1 to 827)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cga@bbs-remail.nih.gov Tissue Procurement: GPCR Consortium cDNA Library Preparation: GPCR Consortium cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: IRB14 row: e column: 10 High quality sequence stop: 653. Location/Qualifiers 1. 827 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30706617" /tissue_type="mixed" /lab_host="DH10B" /clone_lib="NIH_MGC.145" /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XbaI/XhoI-3', 5'-EcoRV-XbaI/NotI-3', EcoRV (7A cloned, non-directional). For information about which gene each clone represents, please visit our anonymous ftp site at ftp://image.lnl.gov/image/rearrayed_plates/IRB1.presv.data a Note: this is a NIH_MGC Library."		

ORIGIN

Alignment Scores:					
Pred. No.:	8,63e-89	Length:	827		
Score:	971.50	Matches:	196		
Percent Similarity:	89.17%	Conservative:	18		
Best Local Similarity:	81.67%	Mismatch:	21		
Query Match:	55.10%	Indels:	5		
Dbl:	7	Gaps:	2		
US-09-867-570-2 (1-337) x CN831885 (1-827)					
OY	100	LennrgrleuLleannlleArghIsProIleser-----lyslle-leUserPProValMe	117		
Dd	785	TTAGGCGCTCAACAAATTATTCAG--CCCATCTCATTCCGCAAAAACCCTCGTTCGTGAT	729		
OY	117	tThr-PheProTyRPhellleglyUSeSerMetleUserrAlalIeserThrsGUAnqCYSL	137		
Dd	728	GACTTTTCCCTACTTTACAGGCCTGAAGTAAGCTGAGGGCCATCAGCACGAGGCGTCC	669		
OY	137	eUserlleleUTRPproIIleTPtyrhlsCyAatgAtgProArTYrlSeSereValM	157		
Dd	668	TGTCTGTTCTTGCGCCCATCTGTGACGCCGCCCCGCCACACCTGTGACGCGTC	609		
OY	157	etCyVsalLeuLeuTriPalaleUserleUeaYgserrlleUeuGUtTrmepheCYsa	177		
Dd	608	TGTGTGTCCTCTCTGGGGCCTGTGTCCTGCTGTTTAGTATGCTGAGTGAAGGTTCTGTG	549		
OY	177	sPpheLeuPheSerglYAlaAspserValTPCYsgIUthrSerAspPheIIleThrllea	197		
Dd	548	ACTTCCTGTATGAGTGNGTCTGATTCGTAGTGTGTGTAACCTGATTCCTCCAGTCG	489		
OY	197	IaTPpleuValPheleuCysValIalleuCYsglySerSerleuValleuValArgI	217		
Dd	488	CGTGCGATTTTTTTATGTGTGTGTCTGTGTGTTTCCAGGCTGTGCTGTCGACAGA	429		
OY	217	IeleUCyglYserArplywmetProleuthrArgleUYrValThrIleuDeuthrv	237		
Dd	428	TCCTCTGTGGATCCCGAAGATGCGCGTGAACAAGGCTGTATGACATCCTGTCTACAG	369		
OY	237	alleuvalPheleuUeucYsglyLeuproPheglylleGIInTPalaLeuPheSerArgI	257		
Dd	368	TGCTGTGTCCTCTCTCGCGGCTGCTGCTTCGCGCATTCGCGGCGCCCTTAATTACAGA	309		
OY	257	IehIsleuAspTRPLyValIleuPheCYshIsValHIsleuValSerIlePheUserrA	277		
Dd	308	TGCACCTGAATTTGGAACTTATAATGTCAATGTATATCTGATTTGCAATGCTCTGCT	249		
OY	277	IaleuAsnSerSerAlaAsnProlleIetyrPhePheValglySerPheArglnaArg	297		
Dd	248	CTCTAAACAGTAGTGCCAACCCCATCATTTACTTCTGTGGGCTCCTTTAGCAGAGCTC	189		
OY	297	IlaenArgrGlAnauleuLytleuValIleuGlnaArgAlaleuGlaAspThrProGUVala	317		
Dd	188	AAAAATAGGCAACCTGAAGCTGTCTCCAGAGGCTCTCAGAGACAAGCTTAAGGTGG	129		
OY	317	spGUlglyGLYLTyrpleuProGlngIunThrleuGUleuSeSerglySerArgrleu	335		
Dd	128	ATTAAGGTGAAGGCGAGCTTCTGTAGAAACCTGTGAGACTGTCCGGAAGCAGATTGG	73		
RESULT 8					
CN839634/C	924 bp	mRNA	linear	EST 02-JUN-2004	
LOCUS	15864260 NIH MGC 145 Homo sapiens cDNA clone				
DEFINITION	IMAGE:30706614 3', mRNA sequence.				
ACCESSION	CN839634				
VERSION	CN839634.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
	1 (bases 1 to 924)				

ORIGIN

/clone_11b="Atherys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is Htt1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in Htt1080 under normal circumstances."

Alignment Scores:

Pred. No.:	2,44e-81	Length:	787
Score:	898.50	Matches:	190
Percent Similarity:	84.34%	Conservative:	20
Best Local Similarity:	76.31%	Mismatches:	39
Query Match:	50.96%	Indels:	4
DB:	4	Gaps:	0

US-09-867-570-2 (1-337) x BG208126 (1-787)

```

Qy 81 LeuAsnLeuValAlaAlaAspPheLeuSerGlyHisIleIleCysSerProLeu 100
Db 747 CTGACGCTGGGGGGGGGACCTTCTCTTTCACAGCCACGTAATCCA-TTCCCTGTTA 689
Qy 101 ArgLeuIleAsnIleArgHisProIleSerIleLeuSerProValMetThrPhePro 120
Db 688 CGCCATCAATAGAGTCATCCCATCTCTTAATCTTATGACCTGATGACCTTTCC 629
Qy 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
Db 628 TACTTGGCAGGCGCTGATATTTCTGAGTGCATAGACACAGGCGCTGCTCAATCTG 569
Qy 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
Db 568 TGGCCATCTGTAAAGCTGCGGCCACCCCA--CACCTGCAACGGTCTGTGTCTG 511
Qy 161 LeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTyrPheCysAspPheLeu 180
Db 510 CTCTGGGCGCTCTCCCTGCTGTAGACATCTCGAGTGTGTTCTGTGACTCCCTGTTT 451
Qy 181 SerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
Db 450 AGTATGCTGATTCGTGTGTGTGCTCAACATCGA-TTCAATCAGATTACGTGCTGATT 392
Qy 201 PheLeuCysValIleLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
Db 391 TTTTATTTGTGTCTCTGTGTGCGCCAGCTGTGTCTGTGTGTGTGTGTGTGTGGA 332
Qy 221 SerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 240
Db 331 TCCGAGAGATCCACTGACAGGCTGTACATGACCAT-CTGCTCACAGTGTGTGCTTTC 273
Qy 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 272 CTCTCTGTGGGCTGCGCCATGTGCATTCAGTGTGCGCTGTTCACAGATCCACATGGAC 213
Qy 261 TrpIleValIlePheCysHisIleValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 212 TGGGAAGCTTATATCTCATGTTCATCTGCTTCCATTTTCTTGGCGTCTCTTAACAGC 153
Qy 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgIleAsnArgGln 300
Db 152 AGTGCACACCCATCTTACTTCTTCATGGGTTTGCTTAGCAGCATCAAAATTAGACAG 93
Qy 301 AsnLeuValLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320
Db 92 AACCTGAAGCTGTTCTCCAGAGGGATCTGCAGAGACACGCTGAGGTGATGAAGGTGA 33
Qy 321 GlyTrpLeuProGlnIleThrLeuGlu 329
Db 32 TGGTGTATCTTCAGGAACCTGTGAG 6

```

RESULT 10
 CN839125

LOCUS CN839125 871 bp mRNA linear EST 02-JUN-2004
 DEFINITION AGENCOURT 15864180 NIH MGC 145 Homo sapiens cDNA clone
 IMAGE:30706609 3', mRNA sequence.
 ACCESSION CN839125
 VERSION CN839125.1 GI:47944780
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 871)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contract: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgaabbs-r@mail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: IRB14 row: e column: 02
 High quality sequence stop: 701.
 Location/Qualifiers

FEATURES

source

1..871
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706609"
 /issue_type="mixed"
 /lab_host="DH10B"
 /clone_11b="NIH MGC 145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
 varies by clone; ORFs were PCR-amplified and cloned into
 pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
 clone and include the following: 5'-EcoRV-XmiI/XhoI-3',
 5'-EcoRV-XmiI/NotI-3', EcoRV (7A cloned, non-directional).
 For information about which gene each clones represents,
 please visit our anonymous ftp site at
 ftp://image.llnl.gov/image/tearrayed_plates/IRB1.presv.dat
 A Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	1.31e-76	Length:	871
Score:	853.00 <td>Matches:</td> <td>172 </td>	Matches:	172
Percent Similarity:	82.43% <td>Conservative:</td> <td>11 </td>	Conservative:	11
Best Local Similarity:	77.48% <td>Mismatches:</td> <td>37 </td>	Mismatches:	37
Query Match:	48.38% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	7 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-09-867-570-2 (1-337) x CN839125 (1-871)

```

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db 195 TCCAGGAGTACACAGACTAGAGGTTTCTGACACATGATCCAAACATCTCAACCTTGACACA 254
Qy 26 GluLeuThrProIleAsnGlyArgGluGluTyrProCysTyrIleValGlnThrLeuSerPhe 45
Db 255 GAACCTACACCAATCAACGAACTGAGGAGCTCTTGTCTACACAGACCTTGAGCCTC 314
Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTrp 65
Db 315 ACGGTGCTACAGTGCATCTTCCCTTGTGCGGCTGACAGGAAAGCAAGTGTGCTGCG 374
Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 375 CTCTGGGCTGCGCATGTGCGAGAACGCTTCTCATCTCATCTCACTTCACTTGGCGCGCA 434

```

QY	86	AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleuIle	105
Db	435	GCAGACTTCCTCTCTCTCAGCGGCGGCTATATATCTCCCTTAAGCTTCATCAATATC	494
QY	106	ArgHisProIleSerIleLeuSerProValMetThrPheProTyRheIleGlyLeu	125
Db	495	CCCCATACCATCTCTAAATCCTCTATCCTGATGATGATTTTCTCACTTCAGGCGCTG	554
QY	126	SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleTyR	145
Db	555	AGCTTTCTGAGTGCCTGAGACCCAGAGCCCTCTGCTCTGCGGCCCATCTGATAC	614
QY	146	HisCysArgArgProArgTyRLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer	165
Db	615	CGCTGCCACGCGCCACACACCTGTGACGGGAGTGATGTGCTCTGCTGGGCGCTGTCC	674
QY	166	LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer	185
Db	675	CTGCTGCGAGAGATCTCGAATGATGATTAATGTGCTTCTGTTCAGTGCTGTGATTTCT	734
QY	186	ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal	205
Db	735	GCTTGGTGTCCAAKATCAGATTTCATACAGACCGCTGCTATATTTTATGTGTGGGT	794
QY	206	LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetP	225
Db	795	CTCTGGGGGCTCAGCGCTGCTCTCTGATTAACGATCTCTGGGGGATTCGCGGAAAAAAC	854
QY	225	ro 225	
Db	855	CC 856	

RESULT 11				
CD051096				
LOCUS				
CD051096	924 bp	mRNA	linear	EST 09-MAY-2003
DEFINITION	AGNSCOURT_1398113	NIH_MGC_173 Homo sapiens	cdna 5'	mRNA sequence

ACCESSION	CD051096
VERSION	CD051096.1
KEYWORDS	GI:30488660
SOURCE	Est.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Buayaotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 924)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
CDNA Library Preparation: Gina Zastrow-Hayes
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: MDKMS5 row: e column: 24
High quality sequence start: 3
High quality sequence stop: 741.
Location/Qualifiers
1. .924

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem
cells"
/lab_host="DH10B Tona"
/clone_lib="Nih MCC173"
/notes="vector: pDONR201; Site 1: attP2; Site 2: attP1;
libR PRIMING - oligo dT; METHOD - full-length enriched;
libR PROVIDER - Bradfield"

```

Alignment Scores:	
Pred. No.:	3,46e-73
Score:	820.00
Percent Similarity:	83.11%
Best Local Similarity:	78.54%
Query Match:	46.51%
DB:	6
Length:	924
Matches:	172
Conservative:	10
Mismatches:	31
Indels:	
Gaps:	0

US-09-867-570-2 (1-337) X CD051096 (1-924

QY 8 ValIleArgLeuGIyPheLeuSerMetAspSerThrIleProValIleGIyThnGIuLeu 27
 Db 245 GTCAACAGACTAGGGGTTTCTGACGATGATCCAAACCATCTCAACCTTGAGACAGAACTG 304
 QY 28 ThrProIleAenGIyArgGIuThnProCysTryIyrsGlnThnLeuSerPheThrGIy 47
 Db 305 ACACCAATCCAAACGAACTGAGAGAACTCTTGTCTACAAAGACAGACTTGAGCCCTACCGGTC 364
 QY 48 LeuThrCysIleValSerLeuValAlaLeuThrGIyAsnAlaValValLeuTrpLeuLeu 67
 Db 365 CTGACGTGCATCGTTTCCCTGTGGGGGTGACAGAAACGGGGTGTGCTCTGGGCTCTG 424
 QY 68 GIyCysArgMetCArgArgAsnAlaValSerIleTryIleLeuAsnLeuValAlaAsp 87
 Db 425 GGCTGCGGCAGCGCCAGGAACGCTTCTCCATCTACATCTCCAACTTGCGCGGACAGAC 484
 QY 88 PheLeuPheLeuSerGIyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHis 107
 Db 485 TTCTCTCTTCCACAGCGCGCCCTATATATATCCCTGGTTAACTTCACATGATCCCCCAT 544
 QY 108 ProIleSerIyIleLeuSerProValMetThrPheProTryPheIleGIyLeuSerMet 127
 Db 545 ACCATCTCTAAATCTCTATCTCTGTGATGATGATTTTCTTACTTGACGGCTGAGCTTT 604
 QY 128 LeuSerAlaIleSerThnGIyArgCysLeuSerIleLeuTrpProIleTrpTryHisCys 147
 Db 605 CTGAGTGGCGGAGACACCGAGCGCTGCGTCCGCGCTGGCCCATCTGGACACCGCTGC 664
 QY 148 ArgArgProArgTryLeuLeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLe 167
 Db 665 CACCGCCCCACACACCTGTGACAGCGGATGTGTGTCTGCTCGGGCCCTGTCCCTGCT 724
 QY 167 uArgSerIleLeuGIuTrpMetPheCysAspPheLeuPheSerGIyAlaAspSerValTr 187
 Db 725 GCGGAGGATCTGTGAGTGCATGATTATGTATGTGCTTCTGTTCAGTGTGCTGATTTCTGCTG 784
 QY 187 pCysGIuThnSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValValLeu 206
 Db 785 GTGTCAAAACATCAATTTTCATCAACAGTCGCGTGGCTGATTTTTTTATGTGTGGGTCTC 844
 QY 207 CysGIySerSerLeuValLeuValArgIleLeuCysGIy 220
 Db 845 TGTGGGTTCAGCGCTGGGTCTGTGGCTGATCAGGAATTTCTGTGGG 889

RESULT 12	871 bp	mRNA	linear	EST 21-APR-2001
LOCUS	BS210740/c			
DEFINITION	RS330287	Athersys RAGE Library	Homo sapiens cDNA, mRNA sequence.	
ACCESSION	BS210740			
VERSION	BS210740.1	GI:13732427		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 871)			
AUTHORS	Harrington,J.J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R., Cain,S., Deventhal,C., Thornton,M., Ramachandran, R., Whittington,J., Letner, L., Costanzo, D., McGillgott, K., Boozer, S., Mays, R., Smith, E., Veloso, N., Kalka, A., Hees, J., Cotthren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.			
TITLE	Creation of genome-wide protein expression libraries using random			

JOURNAL
MEDLINE
11329013
COMMENT
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com

FEATURES
source
High quality sequence step: 456.
Location/Qualifiers
1..871
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_id="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Alignment Scores:
Pred. No.: 1.03e-72 Length: 871
Score: 815.00 Matches: 180
Percent Similarity: 68.14% Conservative: 21
Best Local Similarity: 61.02% Mismatches: 66
Query Match: 46.23% Indels: 29
DB: 4 Gaps: 4

US-09-867-570-2 (1-337) x BG210740 (1-871)

Qy 54 LeuValAlaLeuThrGlyAsnAlaValLeuTrpLeuGlyCysArgMetArgArg 73
Db 867 ATCATAGCTTACTGTTCACACCTACCCGCGTGGCGCGTGGCGGCTCCGCCA 808
Qy 74 AsnAlaValSerIle-----78
Db 807 ACAGCCTGCGCGAGTAAAGTATTCCTTCGTAGACCCCTCATAGGTCGTTCCGTC 748
Qy 79 -----TyrIleLeuAsnLeuValAlaAlaAspPheLeuLeu 91
Db 747 GGGGACACAGCCAGCCCTTCTGTATAGACCGCTATATCCGAGAAATCTTTTCCC 668
Qy 92 SerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerLys 111
Db 667 AGGCCCCGATATATGCGGTAGCTTCG-----AGTTCCCAAAACCTCTCT 640
Qy 112 IleLeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIle 131
Db 639 AAACCTCATCTGTCAGATGTTTCT---ACTTCAGGCGTGTGCTTTTATGTCGTCAG 583
Qy 132 SerThrGlyArgCysLeuSerIleLeuTrpProIleTrpTyrHisCysArgArgProArg 151
Db 582 CACGAAACCCCG---CTGTCCGCTGTGTG---CCCATCTGGTCCGTAGCTACCGCCACCA 527
Qy 152 TyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeu 171
Db 526 CACCTGTGACCGGTGTGTGTCTGTCTGTGGCCCTGTCNTGTGCGGAGACATCTTG 467
Qy 172 GluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSer 191
Db 466 GAGTGATGTATGTGCTTCCCTGTCAGGTGTGCTGATTCCTGTTGTCCTCAACATCA 407
Qy 192 AspPheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeu 211
Db 406 GATTTTCATACACCGCTGCGGTGATTTTATATGTGTGTCTCTGTGGGTCAGGCTG 347
Qy 212 ValLeuLeuValArgIleLeuCysGlySerArgLysMetProLeuThrArgLeuTyrVal 231

Db 346 GTCTCTGATCATGAGATTCTCTGTGATCCGGAGATACCGCTACACGAGCTGACGTG 287
Qy 232 ThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyTlIleTrp 251
Db 286 ACCATCTCTCTCAGTACTGTCTTCTCTCTGTGCGCTGCTTGGCATTCAGTTT 227
Qy 252 AlaLeuPheSerArgIleHisLeuAspTrpLysValLeuPheCysHisValHisLeuVal 271
Db 226 TTCTTATTTTATGATCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 167
Qy 272 SerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGly 291
Db 166 TCATTTCCCTGCTGCGCTCTTAAACAGCAGTGCACCCCATCATTTACTTCTGTCGCGC 107
Qy 292 SerPheArgGlnArgGlnAsnArgGlnIleAsnLeuValLeuValLeuGlnArgAlaLeuGln 311
Db 106 TCTTTAGGACAGCGTCAAAATAGGAGAACCTGAAGCTGTCTTCAGAGGAGGCTTCGAG 47
Qy 312 AspThrProGluValAspGluGlyGlyTyrTrpLeuProGlnGlu 326
Db 46 GACGCGCTGAGCTGATGATGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2

RESULT 13
CN839765/c 953 bp mRNA linear EST 02-JUN-2004
LOCUS
DEFINITION
AGENCOURT 15864292 NIH MGC 145 Homo sapiens cDNA clone
IMAGE:30706616 3', mRNA sequence.
ACCESSION
CN839765
VERSION
CN839765.1 GI:47945420
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9abbs-r@mail.nih.gov
Tissue Procurement: GPCR Consortium
cDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
plate: IRB4 row: e column: 09
High quality sequence step: 578.
Location/Qualifiers
1..953

FEATURES

source

/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706616"
/issue_type="mixed"
/lab_host="DH10B"
/clone_id="NIH MGC 145"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XbaI/XhoI-3', 5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/fearrayed_plates/IRB1.presv.dat a Note: this is a NIH_MGC Library."

ORIGIN

RESULT 15
AK029369
LOCUS
DEFINITION
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4833406120 product:Mrb2 G protein-coupled receptor, full insert sequence.
AK029369
ACCESSION
AK029369.1 GI:26325335
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
6
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Furukawa, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE
Direct Submission
JOURNAL
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-32 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse libraries. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/.

FEATURES
source
1..2075
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4833406120"
/db_xref="taxon:10090"
/clone_type="head"
/clone_id="4833406120"
/clone_idb="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
68..1084
/note="unnamed protein product; Mrb2 G protein-coupled receptor (SPTB) [Q91ZC2, evidence: FASTA, 99.7%id, 100%length, match=1014] putative"
/codon_start=1
/protein_id="BAC26422.1"
/db_xref="GI:26325335"
/translation="MSGDPFLKNTSMAKNTNITVNGSYEDTSVCYTRNQMTLS IITSLVGMGLNVLVLPFGIRNHTNATFVYLLNMAQPLVCSFVCLIAATIFY SIDINIPLVLYVPIFAYLSGLISTISIECLSVIPWYRCRPHHTATCEVL WMSLLGLIEBKACGLFNSPDSYCEFTPDVTNWSVFEVGLSLVLAIFC GSORIPMTLYTITLVLPFLPGLPGIWILOMISNFEYVISCNPYEILFLSC VNSCNPIPIFYVGSIRHRRFRKTLKLLORAMDPTREDSGKSSSEHPELETV QSCS"
2059..2064
/note="putative"
2075
polya_signal
/note="putative"
polya_site
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 8.23e-67 Length: 2075
Score: 763.00 Matches: 159
Percent Similarity: 65.09% Conservative: 48
Best Local Similarity: 50.00% Mismatches: 99
Query Match: 43.28% Indels: 12
DB: 3 Gaps: 5
US-09-867-570-2 (1-337) x AK029369 (1-2075)
QY 13 PheleuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleLeuGly 32
DB 80 TTCCTAATCAAGATTAATTAAGACCTCAGCCTGGAAACACATCATCAGTCTGATGGA 139
QY 33 Arg-----GluGluThrProCysTyrIysGlnThrLeuSerPheThrGlyLeuThr 49
DB 140 AGCTACTACTTGATATCTTCAGTTTGTCACAGGAACCAAGCATGATTTGCTTCC 199
QY 50 CysIleValSerLeuValAlaLeuThrGlyAlaValAlaValLeuTripleLeuGlyCys 69
DB 200 ATCATCATTTCCCTGGTGGATGGACATTAATGCCAATAGTCTGCTGCTGGGCATC 259
QY 70 ArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaSerPheLeu 89
DB 260 GGTATCACACGAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
QY 90 PheLeu---SerGlyHisIleIleCysSerProLeuArgLeu-----Ile 103
DB 320 TACCTGTGCTGCAGTTTATTTGCTTCTTATGCTTCTTATATGCTTCTTACTCAAT 379
QY 104 AsnIleArgHisProIleSerIleLeuSerProValMetThrPheProGlyrHelle 123
DB 380 GACATCAACATCCCTTGG-----GTTCTTATATGTTGGCAATATTTGCTTATCTTCA 433

```
QY 124 GlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIle 143
Db 434 GGCTGAGCAATCTCAGACACCATTAAGCATTAAGCGCTGCTGTGAATATAGGCCATT 493
QY 144 TrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAla 163
Db 494 TGGATAGCGCTGTAACCTCCAGACACACATCAGCTATCAGATTTTGTCCTTGGGTT 553
QY 164 LeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAla 183
Db 554 ATGTCCTTATTTGGGCTCTCCTGGAAGGAGGAGCATGTGGCTTACTGTTTAAATAGCTTT 613
QY 184 AspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCys 203
Db 614 GACTCTTATTTGGTGTGAACATTGATGTATCATATATATGTGCAGTTGTTTTTTT 673
QY 204 ValValLeuCysGlySerSerLeuValIleuValArgIleLeuCysGlySerArgLys 223
Db 674 GGTTCTCTGTGGGCTTAAGCCTCACCCTGCTGTCAAGATCTTCTGTGGCTCACAGCA 733
QY 224 MetProLeuThrArgLeuTyrValThrIleLeuLeuThrValIleuValPheLeuLeuCys 243
Db 734 ATTCTATGACCAAGCTGTATGTAGTACTATTAACAAGCTTGTGCTTCTGATCTTT 793
QY 244 GlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpLysVal 263
Db 794 GGTTCTCCCTTTGGGATCTATTTGATCTATCACTGATTAACAAATTTTATTATGTT 853
QY 264 LeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsn 283
Db 854 GAAATTTGTATTTTATCTTTGAGATACTATCTTATCTCTGTAAACAGCTGTATGAC 913
QY 284 ProIleIleTyrPhePheValGlySerPheArgGlnArgLys--AsnArgGlnAsnLeu 302
Db 914 CCATCATTTATTTCTTGTGGCTCCATTAAGCACCGAAGGTTCAAGCGGAAGACTCTC 973
QY 303 LysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320
Db 974 AAGCTACTTCTGCAGAGGCCATGCAAGACACCCCTGAGGAGAACAAAAGTGA 1027
```

Search completed: December 2, 2004, 01:15:17
Job time : 2986 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:13:59 ; Search time 34384 Seconds
(without alignments)
11858.184 Million cell updates/sec

Title: US-09-867-570-3

Sequence: 1 tgbatgaagcccaatgctcact.....gltccccaagggcccttacc 8622

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_ats: *
12: gb_ey: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8622	100.0	8622	6	AX429467 Sequence
2	8574.6	99.5	91923	9	AC108007 Homo sapi
3	8564.2	99.3	172939	9	AC090099 Homo sapi
4	5966.2	69.2	163718	9	AC023078 Homo sapi
5	5380.4	62.4	165490	2	AC139482 Homo sapi
6	4719	54.7	192618	2	AC020568 Homo sapi
7	4492.2	52.1	192618	2	AC020568 Homo sapi
8	4424.6	51.3	165490	2	AC139482 Homo sapi
9	2601.8	30.2	189230	9	AC103974 Homo sapi
10	1535.8	17.8	156839	9	AC107948 Homo sapi
11	1526.8	17.7	104547	9	AC079120 Homo sapi
12	1521.2	17.6	179216	2	AC074016 Homo sapi
13	1521.2	17.6	181150	9	AC007256 Homo sapi
14	1517.2	17.6	125001	9	AC126613 Homo sapi
15	1507.4	17.5	121520	9	AC006477 Homo sapi
16	1503.4	17.4	155021	9	HS774G10 Human DNA
17	1498.4	17.4	137095	9	AC135452 Homo sapi
18	1497.2	17.4	120591	9	AC116098 Homo sapi
19	1482.6	17.2	102918	9	AC079809 Homo sapi

C 20	1478.2	17.1	205954	9	AC008546 Homo sapi
21	1470	17.0	145250	9	CNS01D18 Homo sapi
22	1462	17.0	74649	9	BX293995 Human DNA
23	1462	17.0	175582	2	AC016977 Homo sapi
24	1462	17.0	321221	2	AL390881 Homo sapi
25	1457	16.9	70466	9	AC068299 Homo sapi
26	1452.8	16.8	193826	9	AL513479 Human DNA
27	1450	16.8	134344	2	HS4773124 Homo sapi
28	1448.4	16.8	211654	2	AL591103 Homo sapi
29	1447.8	16.8	112543	9	AC106714 Homo sapi
30	1447.2	16.8	113428	9	CR626893 Human DNA
31	1447.2	16.8	122867	9	CR626925 Human DNA
32	1447.2	16.8	119858	2	AC020688 Homo sapi
33	1446.8	16.8	110000	2	AL583837_1 Homo sapi
34	1446.8	16.8	135245	9	AL136092 Homo sapi
35	1446.6	16.8	112958	9	AF172277 Homo sapi
36	1445.6	16.8	167358	2	AC010673 Homo sapi
37	1445.6	16.8	178628	2	AC015715 Homo sapi
38	1445.6	16.8	181511	9	AC087738 Homo sapi
39	1445.6	16.8	18937	9	AL360172 Homo sapi
40	1443	16.7	112968	9	AC093833 Homo sapi
41	1442.4	16.7	173460	2	AC138065 Homo sapi
42	1441.2	16.7	173086	2	AL591705 Homo sapi
43	1439.6	16.7	187413	9	AC007567 Homo sapi
44	1438.2	16.7	90526	9	AC092393 Homo sapi
45	1438.2	16.7	133052	9	AL136325 Human DNA

ALIGNMENTS

RESULT 1	AX429467	8622 bp	DNA	linear	PAT 21-JUN-2002
LOCUS	AX429467	Sequence 3 from Patent WO0234914.			
DEFINITION	AX429467				
ACCESSION	AX429467.1	GI:21540764			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	1	Zhao, Q., Beasley, E.M. and Wei, M.H.			
AUTHORS		Isolated human g-protein coupled receptors, nucleic acid molecules encoding human gpcr proteins, and uses thereof			
JOURNAL		Patent: WO 0234914-A 3 02-MAY-2002;			
FEATRES		PR CORP NY (US)			
SOURCE		Location/Qualifiers			
		1..8622			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			
ORIGIN					
Query Match	100.0%	Score 8622;	DB 6;	Length 8622;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 8622;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	TGTTATGAAGCCAAATGTCATTATACCAAAACAGGAAGGATATCAAAAAGAAAC	60		
Db	1	TGTTATGAAGCCAAATGTCATTATACCAAAACAGGAAGGATATCAAAAAGAAAC	60		
Qy	61	TATAGACAGTACCATGATGATATATACATGACAGAAATCCCAAAATACTAGTAC	120		
Db	61	TATAGACAGTACCATGATGATATATACATGACAGAAATCCCAAAATACTAGTAC	120		
Qy	121	CCAAATCAACAGCATATATACAGATATATCCCATTTGTCAGTGGTTTATACCGGG	180		
Db	121	CCAAATCAACAGCATATATACAGATATATCCCATTTGTCAGTGGTTTATACCGGG	180		
Qy	181	GTGAGATAGTGTATATACATGACATGATCAATTAATGTGATCATCATTAAGAAATTA	240		
Db	181	GTGAGATAGTGTATATACATGACATGATCAATTAATGTGATCATCATTAAGAAATTA	240		

Db 181 GTGAGATAGGTAAACATACCAAGTCAATTAATGATGATACATCAATTAACAGATTA 240
 Qy 241 AAAACAAAAATCAATGATCATCTCAATGATGCTGAAAAAGATTGGACAAATCTAAC 300
 Db 241 AAAACAAAAATCAATGATCATCTCAATGATGCTGAAAAAGATTGGACAAATCTAAC 300
 Qy 301 ATTTCTTTATGATTTAAACCTTCAGCAAAATGACATAGAAAAGACATCTTAATGTA 360
 Db 301 ATTTCTTTATGATTTAAACCTTCAGCAAAATGACATAGAAAAGACATCTTAATGTA 360
 Qy 361 TAAAAAGCATATATGACGAGCCACAGCAAACTTAATGATGAGGAAAAAGTTGAAA 420
 Db 361 TAAAAAGCATATATGACGAGCCACAGCAAACTTAATGATGAGGAAAAAGTTGAAA 420
 Qy 421 CATTTCTCTGGAATCTGGAACAGACAGAGATGCTACTTTCACCTTCTATTCAAT 480
 Db 421 CATTTCTCTGGAATCTGGAACAGACAGAGATGCTACTTTCACCTTCTATTCAAT 480
 Qy 481 AGTAGTGAAGTTTATGACGAGCAATCAGACAGAGAAAAGAAATCAAGGACCCAAAT 540
 Db 481 AGTAGTGAAGTTTATGACGAGCAATCAGACAGAGAAAAGAAATCAAGGACCCAAAT 540
 Qy 541 CAATTAAGAGAGATCAAACTGTCTCTGTTCACTGATGATGATGATGATGATGATG 600
 Db 541 CAATTAAGAGAGATCAAACTGTCTCTGTTCACTGATGATGATGATGATGATGATG 600
 Qy 601 CCTTAAGCTGATCCAGAAAGCTCCTAGAACATGATGATGATGATGATGATGATGATG 660
 Db 601 CCTTAAGCTGATCCAGAAAGCTCCTAGAACATGATGATGATGATGATGATGATGATG 660
 Qy 661 ATACAAATCAAAATGATACAAATCAGTACATGCTATACCAACAGTACGACCACTG 720
 Db 661 ATACAAATCAAAATGATACAAATCAGTACATGCTATACCAACAGTACGACCACTG 720
 Qy 721 GAATCAATCAAGAACTCAACACTTTTCAATGATGCTGTAATAAATACTTAAGAAAT 780
 Db 721 GAATCAATCAAGAACTCAACACTTTTCAATGATGCTGTAATAAATACTTAAGAAAT 780
 Qy 781 CTTATCCAGAGAGTGAAGAGCTCTACAAAGAAAATCTACAAACAGTGTGATCATA 840
 Db 781 CTTATCCAGAGAGTGAAGAGCTCTACAAAGAAAATCTACAAACAGTGTGATCATA 840
 Qy 841 GATGACAAACAGAGTGAAGAGCTCTACAAAGAAAATCTACAAACAGTGTGATCATA 900
 Db 841 GATGACAAACAGAGTGAAGAGCTCTACAAAGAAAATCTACAAACAGTGTGATCATA 900
 Qy 901 AAAATGACCAATATGTCAGAAAGCAATCTCAAGTTCAATGCAATCTCCACAAATAT 960
 Db 901 AAAATGACCAATATGTCAGAAAGCAATCTCAAGTTCAATGCAATCTCCACAAATAT 960
 Qy 961 TCATCATCTTCAAGAACTAGAAAATCAATCTTAATAATCTTAATAATCTTAATAAT 1020
 Db 961 TCATCATCTTCAAGAACTAGAAAATCAATCTTAATAATCTTAATAATCTTAATAAT 1020
 Qy 1021 AAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGCTTAGCAAAAAGAAATCAATCTG 1080
 Db 1021 AAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGCTTAGCAAAAAGAAATCAATCTG 1080
 Qy 1081 GGCATCACTTACCCATCTTCAAACTATCTACAGGCTTAATCAACAAACATCATG 1140
 Db 1081 GGCATCACTTACCCATCTTCAAACTATCTACAGGCTTAATCAACAAACATCATG 1140
 Qy 1141 CACTGACATTAAGCACTAGGCACTAGCAATGGAAGAAAGAAAGAAATCCAGAAATTA 1200
 Db 1141 CACTGACATTAAGCACTAGGCACTAGCAATGGAAGAAAGAAAGAAATCCAGAAATTA 1200
 Qy 1201 CCAATTAATTAAGCACTAGGCACTAGGCACTAGGCACTAGGCACTAGGCACTAGG 1260
 Db 1201 CCAATTAATTAAGCACTAGGCACTAGGCACTAGGCACTAGGCACTAGGCACTAGG 1260
 Qy 1261 GACATTTATGTTAAACAAATGCTGATGATTTATGCAAGCCACATGTTGAAATGAAA 1320
 Db 1261 GACATTTATGTTAAACAAATGCTGATGATTTATGCAAGCCACATGTTGAAATGAAA 1320

Qy 1321 CTGATCCCTTGCTCTCACTTAATACAAAATGATACAAAGTATGATCAAAAGCTTAA 1380
 Db 1321 CTGATCCCTTGCTCTCACTTAATACAAAATGATACAAAGTATGATCAAAAGCTTAA 1380
 Qy 1381 TCTGAGACCTAAACCAATTAATTTCTAGAGATTAATCAAGAAATGCTTCTAGAT 1440
 Db 1381 TCTGAGACCTAAACCAATTAATTTCTAGAGATTAATCAAGAAATGCTTCTAGAT 1440
 Qy 1441 TCACTTAGGCAAGACTTATGAGGCAAGAACCCAAAGTAAATGCAACAAAAT 1500
 Db 1441 TCACTTAGGCAAGACTTATGAGGCAAGAACCCAAAGTAAATGCAACAAAAT 1500
 Qy 1501 AAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
 Db 1501 AAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
 Qy 1561 AGCAAAACAGAACCAACCAAGTGAAGAAATCTTCAAAATCAAGATCTGACTAAG 1620
 Db 1561 AGCAAAACAGAACCAACCAAGTGAAGAAATCTTCAAAATCAAGATCTGACTAAG 1620
 Qy 1621 ACTAATATCCGAAATCCAGAGAACTCAACAAATCAAGCAAGAAAGAAAGCAATCC 1680
 Db 1621 ACTAATATCCGAAATCCAGAGAACTCAACAAATCAAGCAAGAAAGAAAGCAATCC 1680
 Qy 1681 CATGAAGAGTGGGCTAAGAGATGATGATGATGATGATGATGATGATGATGATG 1740
 Db 1681 CATGAAGAGTGGGCTAAGAGATGATGATGATGATGATGATGATGATGATGATG 1740
 Qy 1741 CAACAAACAGAAATAATGCTTAATCAATCAATGATGATGATGATGATGATGATG 1800
 Db 1741 CAACAAACAGAAATAATGCTTAATCAATCAATGATGATGATGATGATGATGATG 1800
 Qy 1801 TAATGATATCACTTACTCTGCAAGATGCTATTAATTAATTAATTAATTAATTA 1860
 Db 1801 TAATGATATCACTTACTCTGCAAGATGCTATTAATTAATTAATTAATTAATTA 1860
 Qy 1861 GATGTTGGGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 Db 1861 GATGTTGGGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 Qy 1921 TTGCGCAACCAATATGAAAAAGTGTGAAATTTCTTAAGAACTAAAGTATGATG 1980
 Db 1921 TTGCGCAACCAATATGAAAAAGTGTGAAATTTCTTAAGAACTAAAGTATGATG 1980
 Qy 1981 CATTTGATCAAGATCCCATTAATTAATGATTAATTAATTAATTAATTAATTAAT 2040
 Db 1981 CATTTGATCAAGATCCCATTAATTAATGATTAATTAATTAATTAATTAATTAAT 2040
 Qy 2041 ACAACTCAGCCTAATAAAGAAATTAATGATGATGATGATGATGATGATGATGATG 2100
 Db 2041 ACAACTCAGCCTAATAAAGAAATTAATGATGATGATGATGATGATGATGATGATG 2100
 Qy 2101 AGACCTTAATTTAAGTGGGTAATCTAGAAATGAAAAACCAATCATATATGTTCTAC 2160
 Db 2101 AGACCTTAATTTAAGTGGGTAATCTAGAAATGAAAAACCAATCATATATGTTCTAC 2160
 Qy 2161 TTAACAAGTGGGGCTTAAGTGTGAGAGACAGAAAGCAATGATGATTAATTAATG 2220
 Db 2161 TTAACAAGTGGGGCTTAAGTGTGAGAGACAGAAAGCAATGATGATTAATTAATG 2220
 Qy 2221 GGACTTGAAGGGAAGATGGAAGAGAGGAGGAGGATTAAGATCAACATGAGGATCAG 2280
 Db 2221 GGACTTGAAGGGAAGATGGAAGAGAGGAGGAGGATTAAGATCAACATGAGGATCAG 2280
 Qy 2281 GTACATGCTCAGGTGATGAGTGAACAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340
 Db 2281 GTACATGCTCAGGTGATGAGTGAACAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340
 Qy 2341 CATGGAAGCAACCACTGTTCCCAAAATCCCAATGAAATTAATTAATTAATTAATTA 2400
 Db 2341 CATGGAAGCAACCACTGTTCCCAAAATCCCAATGAAATTAATTAATTAATTAATTA 2400

QY	2401	TTGATTTAAATTTCACAGAAATTTAAAAAGTTCACTGTTCAAGGTTTAATAATAAGAACTA	2460
Db	2401	ATGATTTAAATTTCACAGAAATTTAAAAAGTTCACTGTTCAAGGTTTAATAATAAGAACTA	2460
QY	2461	AGAAATGAAGAGTGTGCAAGTGTGTGCTCTGAGCAATAGGAGCTCTAATTTTCAACTGT	2520
Db	2461	AGAAATGAAGAGTGTGCAAGTGTGTGCTCTGAGCAATAGGAGCTCTAATTTTCAACTGT	2520
QY	2521	CATACACTTCTCTGGCATTTTGGAAAGAAAGTATACATGAAATATATACACCATATGATA	2580
Db	2521	CATACACTTCTCTGGCATTTTGGAAAGAAAGTATACATGAAATATATACACCATATGATA	2580
QY	2581	AAGAAAAACATCAAAAAATTTGTGTCAAGGCCATTTGTCAAGCTTGAATGTGTCCATGATCTAC	2640
Db	2581	AAGAAAAACATCAAAAAATTTGTGTCAAGGCCATTTGTCAAGCTTGAATGTGTCCATGATCTAC	2640
QY	2641	TTTTTCATTTTGGATATAAAGCCTCATATATGATGTTCATTTGCTTAATATGATAGCCCTAG	2700
Db	2641	TTTTTCATTTTGGATATAAAGCCTCATATATGATGTTCATTTGCTTAATATGATAGCCCTAG	2700
QY	2701	GCCCATAAATTGATTTTTTAAAAATCAGGACAGCAATTCTTAACAGAAAGTTGAAACAAGATGG	2760
Db	2701	GCCCATAAATTGATTTTTTAAAAATCAGGACAGCAATTCTTAACAGAAAGTTGAAACAAGATGG	2760
QY	2761	GACGTGATATAGAGAGGCTTAAATGTATCTGATATGGAACAAGGCCAAGAAATCATCTCAG	2820
Db	2761	GACGTGATATAGAGAGGCTTAAATGTATCTGATATGGAACAAGGCCAAGAAATCATCTCAG	2820
QY	2821	TTAGGATTTGTGTCTCAAAATCCTCTGGCCCTGTGATTTGGCCATATGTCTCATACAGAA	2880
Db	2821	TTAGGATTTGTGTCTCAAAATCCTCTGGCCCTGTGATTTGGCCATATGTCTCATACAGAA	2880
QY	2881	ATTAACAAGACTGTCCAGACATCTTCTGAAGCCTGATTTCTCAACAGCTTTCATTTCAAGCT	2940
Db	2881	ATTAACAAGACTGTCCAGACATCTTCTGAAGCCTGATTTCTCAACAGCTTTCATTTCAAGCT	2940
QY	2941	CCTGTAGGCATCTCTGAAATTAAGCAACAAGAAAAGTCTCTGAAAGTCACTGAAATCCCA	3000
Db	2941	CCTGTAGGCATCTCTGAAATTAAGCAACAAGAAAAGTCTCTGAAAGTCACTGAAATCCCA	3000
QY	3001	GAAAGGCTCTCACTCTTTAGCAACAAGGAGGCTTCAACACTGACAAAGAAAGAAACAT	3060
Db	3001	GAAAGGCTCTCACTCTTTAGCAACAAGGAGGCTTCAACACTGACAAAGAAAGAAACAT	3060
QY	3061	AAGGTAATGATACAAAGAACTCTCTCTTCCACAGTCAGATTATGATTTTGTGCTGTAAAGTC	3120
Db	3061	AAGGTAATGATACAAAGAACTCTCTCTTCCACAGTCAGATTATGATTTTGTGCTGTAAAGTC	3120
QY	3121	ATGTCCTTATGCTTTCACACTTGTGTGCTACATGCAAGGGGCTACGAGCTTGTTCAGGAAA	3180
Db	3121	ATGTCCTTATGCTTTCACACTTGTGTGCTACATGCAAGGGGCTACGAGCTTGTTCAGGAAA	3180
QY	3181	AGACAGGAGACATGAAGCTTCTTTCAAGAACTGAGTGTCTCAACCCAAACTGTGTGAG	3240
Db	3181	AGACAGGAGACATGAAGCTTCTTTCAAGAACTGAGTGTCTCAACCCAAACTGTGTGAG	3240
QY	3241	CTCTAAATAGTGTGCCCCCTTCTAATTTATCTCCCAATATCACTCTCTTCAATTCATATCA	3300
Db	3241	CTCTAAATAGTGTGCCCCCTTCTAATTTATCTCCCAATATCACTCTCTTCAATTCATATCA	3300
QY	3301	TTCAATCTGCTCATGAGAGACTGTGCTCTTACATTCATTAACATTTAAAGACAAAGGGAC	3360
Db	3301	TTCAATCTGCTCATGAGAGACTGTGCTCTTACATTCATTAACATTTAAAGACAAAGGGAC	3360
QY	3361	ATGCAAGGCAATTTCTTCCAGAGTTGAACCTGCTATAGAGCCAGTTTCTTTGTTTCACTTAC	3420
Db	3361	ATGCAAGGCAATTTCTTCCAGAGTTGAACCTGCTATAGAGCCAGTTTCTTTGTTTCACTTAC	3420
QY	3421	TTTTCAAAATTAATTTCTTCTTTGGCTATCTGAAAGGCTTAAGAAAGATATATGATAGGCCCA	3480
Db	3421	TTTTCAAAATTAATTTCTTCTTTGGCTATCTGAAAGGCTTAAGAAAGATATATGATAGGCCCA	3480
QY	3481	ATAAATTAAGAGGTGTTTCATGAGAAAGTATTTAACAAAGATGCAACAAGTTAAGGGCTAG	3540

Db	3481	ATAATTAAGGAGTGTTCATAGAGAAATATTTCAAGATGCAACAGTTAAAGGCTAG	3540
Oy	3541	GATCCTTAAGCACAATACATAGAGGAGACCTACTTCCCTCCCTAAGGCTGAACGACAGG	3600
Db	3541	GATCCTTAAGCACAATACATAGAGGAGACCTACTTCCCTCCCTAAGGCTGAACGACAGG	3600
Oy	3601	GAAGGAGCATTTACCATTTGTGCGCATAGCCATAGCTGTAGCCATAGGGTGGAGAGCAT	3660
Db	3601	GAAGGAGCATTTACCATTTGTGCGCATAGCCATAGCTGTAGCCATAGGGTGGAGAGCAT	3660
Oy	3661	GAGCAGGCAACTGAGGAAGCCCTGCGTGGCCAAAGCACAAGCCACAAGGCTGATATAGTT	3720
Db	3661	GAGCAGGCAACTGAGGAAGCCCTGCGTGGCCAAAGCACAAGCCACAAGGCTGATATAGTT	3720
Oy	3721	TGATCTGTGTGCCACCAAAATCATAGTTGATTTGTAATTTCCAAATGTTGAGAGAAAGG	3780
Db	3721	TGATCTGTGTGCCACCAAAATCATAGTTGATTTGTAATTTCCAAATGTTGAGAGAAAGG	3780
Oy	3781	CCTTGTGGAGATGATTTATAGATCACGGGGATGGTTTTGCATGATGTTTTAAACCAT	3840
Db	3781	CCTTGTGGAGATGATTTATAGATCACGGGGATGGTTTTGCATGATGTTTTAAACCAT	3840
Oy	3841	CCCCCTTGGTATTTGTTGTGTGTGATCTGACAGAGTTCTCATGAAATCTAGTTGTTTAAA	3900
Db	3841	CCCCCTTGGTATTTGTTGTGTGTGATCTGACAGAGTTCTCATGAAATCTAGTTGTTTAAA	3900
Oy	3901	GGGTGTAGCACTCCCTCCCTCCTTATCTCCTGTCTCACCATGTGAGAGCCCTCGCTC	3960
Db	3901	GGGTGTAGCACTCCCTCCCTCCTTATCTCCTGTCTCACCATGTGAGAGCCCTCGCTC	3960
Oy	3961	CCCCCTTGGCCTTTCACAGAGATTGGAAGCTTCTGAGGCTCTCCACAGAGCAGAAGCTGC	4020
Db	3961	CCCCCTTGGCCTTTCACAGAGATTGGAAGCTTCTGAGGCTCTCCACAGAGCAGAAGCTGC	4020
Oy	4021	TATGCTTCTTGTACAGTCTGTAGAGCTATTAGCCAGTTAAACCATTTCCCTTCATAAAT	4080
Db	4021	TATGCTTCTTGTACAGTCTGTAGAGCTATTAGCCAGTTAAACCATTTCCCTTCATAAAT	4080
Oy	4081	TCCCACTCTCAGGATTTCTTTTATAGCAATTTGGAATGAATCAATACACAGACAGAG	4140
Db	4081	TCCCACTCTCAGGATTTCTTTTATAGCAATTTGGAATGAATCAATACACAGACAGAG	4140
Oy	4141	CCAGGAGATGGAATCCCAAGGTGCTTCTGCTGTCTTCCAGTCTCTGCTGGTGTCTC	4200
Db	4141	CCAGGAGATGGAATCCCAAGGTGCTTCTGCTGTCTTCCAGTCTCTGCTGGTGTCTC	4200
Oy	4201	CCAGTGTCTCAATTCCACAGAAAACAGAAATYAAAAAGATCCCATGTGTGTACATA	4260
Db	4201	CCAGTGTCTCAATTCCACAGAAAACAGAAATYAAAAAGATCCCATGTGTGTACATA	4260
Oy	4261	GAAGGCACTCTTGGGATGTCAAAACAGGATTAAGAAAGAAAGAAATCTCATGG	4320
Db	4261	GAAGGCACTCTTGGGATGTCAAAACAGGATTAAGAAAGAAAGAAATCTCATGG	4320
Oy	4321	TAAATGACATATCCCTCTCACTTCTGTATCTCCTCAATTCCTGGGGCTTCTCATATC	4380
Db	4321	TAAATGACATATCCCTCTCACTTCTGTATCTCCTCAATTCCTGGGGCTTCTCATATC	4380
Oy	4381	TGATTTGATCCCTGTCTCAATTCACGCTCATACATCTTAAATGTTTGGCTTGTCTTC	4440
Db	4381	TGATTTGATCCCTGTCTCAATTCACGCTCATACATCTTAAATGTTTGGCTTGTCTTC	4440
Oy	4441	TCTACTGTCACTTTATGCAAGAAATGTTGATTTGTAATAAATGCAATAGAAATPAAAT	4500
Db	4441	TCTACTGTCACTTTATGCAAGAAATGTTGATTTGTAATAAATGCAATAGAAATPAAAT	4500
Oy	4501	GTAATTTTAAAAAGACATATGATATTTTGTATGAATATATAGTTTGGCTGATCTAATAA	4560
Db	4501	GTAATTTTAAAAAGACATATGATATTTTGTATGAATATATAGTTTGGCTGATCTAATAA	4560
Oy	4561	GACATGAAGAAATATCTTAAACAGAAAGTATAGTTGTGCTTGGCTCACTAGGTT	4620
Db	4561	GACATGAAGAAATATCTTAAACAGAAAGTATAGTTGTGCTTGGCTCACTAGGTT	4620

Db	4561	GACATGAAGAAAGAAATCTTAAACAAGAAAGTATAGTTGGCCCTCGGGCTACAGATT	4620
OY	4621	CTGAATCTACAGATTCAACAACCTACAGAGAGAACTTTTCCAAAATTAAGTGTGCG	4680
Db	4621	CTGAATCTACAGATTCAACAACCTACAGAGAGAACTTTTCCAAAATTAAGTGTGCG	4680
OY	4661	GAGTTGTGTACTGTGAACAGGTACAAACTTGATTTCTTGTCATATATTTCTGAAAA	4740
Db	4661	GAGTTGTGTACTGTGAACAGGTACAAACTTGATTTCTTGTCATATATTTCTGAAAA	4740
OY	4741	CTACAAATTAACAAGAACTTATATATAGACTTTTGTCAGTTATCTAAATACCTT	4800
Db	4741	CTACAAATTAACAAGAACTTATATATAGACTTTTGTCAGTTATCTAAATACCTT	4800
OY	4801	AAATGATTTAATGATCTGGAGAAAGTGCATAGATATATACAAATACATATATAGG	4860
Db	4801	AAATGATTTAATGATCTGGAGAAAGTGCATAGATATATACAAATACATATATAGG	4860
OY	4861	AAATTGAGACTGTGCAGATTTTGGTCTGTGCTGGGGTCTGGAAGAAATCCCTGTAA	4920
Db	4861	AAATTGAGACTGTGCAGATTTTGGTCTGTGCTGGGGTCTGGAAGAAATCCCTGTAA	4920
OY	4921	ACACAAAATGACACTCTTCGAGCTGAACTAGAAAGCTCCAAAGCATCATATACGAA	4980
Db	4921	ACACAAAATGACACTCTTCGAGCTGAACTAGAAAGCTCCAAAGCATCATATACGAA	4980
OY	4981	TTCCAAAAATTGCTGCTCCCGCAGTTCCTAGAGAGTTGCGCTCATCTTGATCTCAT	5040
Db	4981	TTCCAAAAATTGCTGCTCCCGCAGTTCCTAGAGAGTTGCGCTCATCTTGATCTCAT	5040
OY	5041	GGTTCCTCCAGCGACATTAGACTTCCAGTCTTATGAAAAAGACGAGGGGAGAGAGCT	5100
Db	5041	GGTTCCTCCAGCGACATTAGACTTCCAGTCTTATGAAAAAGACGAGGGGAGAGAGCT	5100
OY	5101	TTGCTCCTTATTAATGATCCATGAGACGAGACTTGCTCTGACCTTTTGATCTTCC	5160
Db	5101	TTGCTCCTTATTAATGATCCATGAGACGAGACTTGCTCTGACCTTTTGATCTTCC	5160
OY	5161	ACTTAACAGACACTCTGCTCATGAGGATGTATCCAGCATCAAGAAAACTGGATGTGGTC	5220
Db	5161	ACTTAACAGACACTCTGCTCATGAGGATGTATCCAGCATCAAGAAAACTGGATGTGGTC	5220
OY	5221	CTTGCTGCTGTGATCATCTTCAGAAAGGTATGTGACCAAAAAAGGAACTTGCGGCA	5280
Db	5221	CTTGCTGCTGTGATCATCTTCAGAAAGGTATGTGACCAAAAAAGGAACTTGCGGCA	5280
OY	5281	ACCACACAGCTCTTAGGCCCTGACCGTCTCGATTTCTGTGCTCATCAAGATTTTCA	5340
Db	5281	ACCACACAGCTCTTAGGCCCTGACCGTCTCGATTTCTGTGCTCATCAAGATTTTCA	5340
OY	5341	GGAATCTCTCAGAAATTAATTAATGATGGGGCAGAGAACAGAACTGATCTCTGACAGA	5400
Db	5341	GGAATCTCTCAGAAATTAATTAATGATGGGGCAGAGAACAGAACTGATCTCTGACAGA	5400
OY	5401	CTTCAGGAGACGAGGGCTGTGATTTGAACTGTCTTCATGTGTGAAACAAGAAAAACCT	5460
Db	5401	CTTCAGGAGACGAGGGCTGTGATTTGAACTGTCTTCATGTGTGAAACAAGAAAAACCT	5460
OY	5461	TTAATTTCTTAGGCTTAGCTTCATCTTATGTATATAGAGATAATACATAGACGCT	5520
Db	5461	TTAATTTCTTAGGCTTAGCTTCATCTTATGTATATAGAGATAATACATAGACGCT	5520
OY	5521	TTAAAGAACATCATAGACATGTAAACAACATCTTAAATGTTGGTATACCAAGTGAAGA	5580
Db	5521	TTAAAGAACATCATAGACATGTAAACAACATCTTAAATGTTGGTATACCAAGTGAAGA	5580
OY	5581	AGACAGGACATGACTTACTCTTACGGATCTTGGGGTTTCAAGAGAAACAACATATCA	5640
Db	5581	AGACAGGACATGACTTACTCTTACGGATCTTGGGGTTTCAAGAGAAACAACATATCA	5640
OY	5641	TACCATCTCTATAGATGCAAAACAGTTTATAGTGTCTGAGTGTGATTAACAGAGTTCTC	5700
Db	5641	TACCATCTCTATAGATGCAAAACAGTTTATAGTGTCTGAGTGTGATTAACAGAGTTCTC	5700

QY	5701	TTTTTCCCTCCCAATTTCCCTTTTGGGCGCAATCAGAGCTGAGAGCTGCTCCCTTAAG	5760
Db	5701	CTTTTCCCTCCCAATTTCCCTTTTGGGCGCAATCAGAGCTGAGAGCTGCTCCCTTAAG	5760
QY	5761	AGCTCATGATGATGACACTCACTCTGATGCTCCTCATATCTCCAGAGAGATGATC	5820
Db	5761	AGCTCATGATGATGACACTCACTCTGATGCTCCTCATATCTCCAGAGAGATGATC	5820
QY	5821	TTCTTTCCACTGAGAGCTCTGCGCATGTGCATTTTGGGATTTCCAGGCAACGTGG	5880
Db	5821	TTCTTTCCACTGAGAGCTCTGCGCATGTGCATTTTGGGATTTCCAGGCAACGTGG	5880
QY	5881	CCTGATAGGCAAAAAAGAACTCCTGAATTTGTTTCTTAATGGCAGCACTACCTCTA	5940
Db	5881	CCTGATAGGCAAAAAAGAACTCCTGAATTTGTTTCTTAATGGCAGCACTACCTCTA	5940
QY	5941	TTTTTCCCTTAATTCATTTGCTTCTCATTTCTCATCTGAGATTGTTTAAAGTTAATTTT	6000
Db	5941	TTTTTCCCTTAATTCATTTGCTTCTCATTTCTCATCTGAGATTGTTTAAAGTTAATTTT	6000
QY	6001	TTTTTACGCCCAATTTTGACTGTGCACTTGAATTAACTGAGATCACTCCTCACT	6060
Db	6001	TTTTTACGCCCAATTTTGACTGTGCACTTGAATTAACTGAGATCACTCCTCACT	6060
QY	6061	TTACCCCTCTTAACATGTATATACGACATAGTGTGTGGTCCMAAGGGCTGTGA	6120
Db	6061	TTACCCCTCTTAACATGTATATACGACATAGTGTGTGGTCCMAAGGGCTGTGA	6120
QY	6121	AAAAATGATCATAGTACAGCCCTGCTGGCTCACTTCACTATATATATATACC	6180
Db	6121	AAAAATGATCATAGTACAGCCCTGCTGGCTCACTTCACTATATATATATACC	6180
QY	6181	CCCGGACAAATATATCCTCTCTTATACTTATATTTCAATATCGAATACAGAAATA	6240
Db	6181	CCCGGACAAATATATCCTCTCTTATACTTATATTTCAATATCGAATATACAGAAATA	6240
QY	6241	TACTAATTTTACCTCCTAGGCTCTTGAGATTAATAAGGCAATCCTAATTAACCTG	6300
Db	6241	TACTAATTTTACCTCCTAGGCTCTTGAGATTAATAAGGCAATCCTAATTAACCTG	6300
QY	6301	TCATACAGCTGTATATCTCCCAAAATTTAGACTATTCCTCATTCCTCCAGTTGAATTT	6360
Db	6301	TCATACAGCTGTATATCTCCCAAAATTTAGACTATTCCTCATTCCTCCAGTTGAATTT	6360
QY	6361	GCATGAATATCTCTTTACACCACCAAGCCCTACACTTCTCTAATTTCACTCACTAGACT	6420
Db	6361	GCATGAATATCTCTTTACACCACCAAGCCCTACACTTCTCTAATTTCACTCACTAGACT	6420
QY	6421	CCTCTCATACAAATGTTTGATCAACAAGAAAGCTATCCAAAGATCTCCCGAAAGAG	6480
Db	6421	CCTCTCATACAAATGTTTGATCAACAAGAAAGCTATCCAAAGATCTCCCGAAAGAG	6480
QY	6481	AATGAAATAGGTTTACATTTGTGTATCTCAGCAGAAACACTTAGTAGTCCCCATACATAT	6540
Db	6481	AATGAAATAGGTTTACATTTGTGTATCTCAGCAGAAACACTTAGTAGTCCCCATACATAT	6540
QY	6541	TTCCACACTTCAATTAACCTGTGCAAGTGGCACTCAGGCTCACCTCACTTACTTCTTCT	6600
Db	6541	TTCCACACTTCAATTAACCTGTGCAAGTGGCACTCAGGCTCACCTCACTTACTTCTTCT	6600
QY	6601	CTGTTCTATGCTAGCAATTCAGCTCAGACCCACACCTTACCMAACATGTGTACAA	6660
Db	6601	CTGTTCTATGCTAGCAATTCAGCTCAGACCCACACCTTACCMAACATGTGTACAA	6660
QY	6661	ATGCTTTAGGGGTTCCGCAAGGCAACCTGAGTCTTATTTTAAAGCACTACGTGT	6720
Db	6661	ATGCTTTAGGGGTTCCGCAAGGCAACCTGAGTCTTATTTTAAAGCACTACGTGT	6720
QY	6721	CAATTTACAGTTTGGGCACTCATCATCATCTTCTTCAACACAGATAGAGCTGTCCACA	6780
Db	6721	CAATTTACAGTTTGGGCACTCATCATCATCTTCTTCAACACAGATAGAGCTGTCCACA	6780

QY	6781	AATGAATTCGATGAAGAAATTTTCTCATCTAAATTAATAGTGTCCTAAATGCT	6840
Db	6781	AATGAATTCGATGAAGAAATTTTCTCATCTAAATTAATAGTGTCCTAAATGCT	6840
QY	6841	TACATTGTCCTTCAATTTTATTTTCCATTCACAATCTACCATTTGCCATTAGGCTT	6900
Db	6841	TACATTGTCCTTCAATTTTATTTTCCATTCACAATCTACCATTTGCCATTAGGCTT	6900
QY	6901	CTCATGCAATGATTCCTTCATTGAATGAACGTTTATGAAGACAACTGTGCTGTTATG	6960
Db	6901	CTCATGCAATGATTCCTTCATTGAATGAACGTTTATGAAGACAACTGTGCTGTTATG	6960
QY	6961	GAATNAGGCACTNAGAGATTAATAATGTAAATGTGTGCTGTCTGCAATGATGACACT	7020
Db	6961	GAATNAGGCACTNAGAGATTAATAATGTAAATGTGTGCTGTCTGCAATGATGACACT	7020
QY	7021	GAGTTATTTCTCACCCACAGAGTCCCGCATTTTTCACACATCTAGCGAAATCCCATTT	7080
Db	7021	GAGTTATTTCTCACCCACAGAGTCCCGCATTTTTCACACATCTAGCGAAATCCCATTT	7080
QY	7081	TCCTCTGGTTCAATAAGCATATCTTTTTCCTGTCCAGAGATGACCAAGTCTGTCATG	7140
Db	7081	TCCTCTGGTTCAATAAGCATATCTTTTTCCTGTCCAGAGATGACCAAGTCTGTCATG	7140
QY	7141	AGGGTGTCAACAACCCCTTTTGTGTATCTGAATTCCTCCACCTGAGAGAAATTTTCAG	7200
Db	7141	AGGGTGTCAACAACCCCTTTTGTGTATCTGAATTCCTCCACCTGAGAGAAATTTTCAG	7200
QY	7201	CCCAAGATNAGATATCATCGGGGTCCACAGACACTGGCTAGATNAGTGGGGGTGTTTAT	7260
Db	7201	CCCAAGATNAGATATCATCGGGGTCCACAGACACTGGCTAGATNAGTGGGGGTGTTTAT	7260
QY	7261	CCTAATGTTATCCCATGTGTCAGACAGAACTTGTGTGCAATGAGAGAGGTCAAGCTTC	7320
Db	7261	CCTAATGTTATCCCATGTGTCAGACAGAACTTGTGTGCAATGAGAGAGGTCAAGCTTC	7320
QY	7321	AGAGTCAACAAGAACTGATTTCAAACTGTGATTTGAGGACCCCACTTTTGAATAGTGA	7380
Db	7321	AGAGTCAACAAGAACTGATTTCAAACTGTGATTTGAGGACCCCACTTTTGAATAGTGA	7380
QY	7381	CTTATTTCTGTGCGAGTCTGTATCTCTCTCTTAAATGAGGACAGTAAATCCCAATGG	7440
Db	7381	CTTATTTCTGTGCGAGTCTGTATCTCTCTCTTAAATGAGGACAGTAAATCCCAATGG	7440
QY	7441	CAGGGTGTGGGGAGAACTCAGAGATCAAAACAGCTGTGATCAATCTGGTTCGTGTTCC	7500
Db	7441	CAGGGTGTGGGGAGAACTCAGAGATCAAAACAGCTGTGATCAATCTGGTTCGTGTTCC	7500
QY	7501	AGGGTCAATCAGACTGGGGTTTCTGAGCATGAGATTCACCATCCAGTCTTGGGTACAGA	7560
Db	7501	AGGGTCAATCAGACTGGGGTTTCTGAGCATGAGATTCACCATCCAGTCTTGGGTACAGA	7560
QY	7561	CTGACACCAATCAACGGAAGGAGTGAAGGACTCTCTTGCTAAGGACAGCCCTGAGCTTACG	7620
Db	7561	CTGACACCAATCAACGGAAGGAGTGAAGGACTCTCTTGCTAAGGACAGCCCTGAGCTTACG	7620
QY	7621	GGGCTGACGTGATCTGTTTCCCTTGTGGCGCTGACAGGAAACGGGTTGTCTCTGGCTC	7680
Db	7621	GGGCTGACGTGATCTGTTTCCCTTGTGGCGCTGACAGGAAACGGGTTGTCTCTGGCTC	7680
QY	7681	CTGGGCTGCGGCATGCGGAGAAAGCTGTCTTCATCTAATCTTCAACCTGAGTGGCGGC	7740
Db	7681	CTGGGCTGCGGCATGCGGAGAAAGCTGTCTTCATCTAATCTTCAACCTGAGTGGCGGC	7740
QY	7741	GACTTCCCTTCTCTTAAAGGCGCACTAATATGTTTCGCGTTACGCTCATCAATATCCGC	7800
Db	7741	GACTTCCCTTCTCTTAAAGGCGCACTAATATGTTTCGCGTTACGCTCATCAATATCCGC	7800
QY	7801	CATCCCATCTCCAAATCTCAAGTCTGTGATGACCTTCCCTAATTATAGGCGCTAAGC	7860
Db	7801	CATCCCATCTCCAAATCTCAAGTCTGTGATGACCTTCCCTAATTATAGGCGCTAAGC	7860
QY	7861	ATGCTGAGCGCCATCAGACCGAGCGTGCCTGTTCATCTGTGGCCATCTGTATCAC	7920

Db	7861	ATGCTGAGCCGCATCAGCACCGAGCCGCTGCCTGTCATCCGTGCGCCATCTGGTACAC	7920
Qy	7921	TGCCGCGCCGCCAGATACCTGTCAATGCGATCATGTGTCTCTGCTGCGCCCTGTCCCTG	7980
Db	7921	TGCCGCGCGCCGCCAGATACCTGTCAATGCGATCATGTGTCTCTGCTGCGCCCTGTCCCTG	7980
Qy	7981	CTGGGAGATATCTCGAGAGTGAATGTTCTGTGACCTTCTGTTTAAATGAGTGCATATTCCTT	8040
Db	7981	CTGGGAGATATCTCGAGAGTGAATGTTCTGTGACCTTCTGTTTAAATGAGTGCATATTCCTT	8040
Qy	8041	TGAGTGAAGAGTACGATTTTCATTAACATCCCGAGGCTGGTTTTTTTAAATGAGTGCCTC	8100
Db	8041	TGAGTGAAGAGTACGATTTTCATTAACATCCCGAGGCTGGTTTTTTTAAATGAGTGCCTC	8100
Qy	8101	TGTGGGCTCCAGCCCTGATCTGCTGCTGATGAGATTTCTCTGTGTGATCCCGAAGATGCGCTG	8160
Db	8101	TGTGGGCTCCAGCCCTGATCTGCTGCTGATGAGATTTCTCTGTGTGATCCCGAAGATGCGCTG	8160
Qy	8161	ACCAAGGCTGATCGATGACCATCTCTCTCAAGTGTGCTCTTCTCTCTGTGTGCTGCTGCC	8220
Db	8161	ACCAAGGCTGATCGATGACCATCTCTCTCAAGTGTGCTCTTCTCTCTGTGTGCTGCTGCC	8220
Qy	8221	TTTGGCATTCAGTGGGCGCTGTTTTCAGAGATTCACCGATGATGAGAAAGTCTTAATTTTGT	8280
Db	8221	TTTGGCATTCAGTGGGCGCTGTTTTCAGAGATTCACCGATGATGAGAAAGTCTTAATTTTGT	8280
Qy	8281	CATGTGATCATGATTTTCATTTTCTGTCGCTCTTAAACAGAGTGCACCCCATCATTT	8340
Db	8281	CATGTGATCATGATTTTCATTTTCTGTCGCTCTTAAACAGAGTGCACCCCATCATTT	8340
Qy	8341	TACTTCTTCTGTGGCTCTTTTAAAGAGGCTCAAAATAGCGAAGCTGAGCTGCTTC	8400
Db	8341	TACTTCTTCTGTGGCTCTTTTAAAGAGGCTCAAAATAGCGAAGCTGAGCTGCTTC	8400
Qy	8401	CAGAGGGCTCTGACAGAACGCGCTGAGGTGATTAAGGTGAGGGTGGCTTCCACAGAA	8460
Db	8401	CAGAGGGCTCTGACAGAACGCGCTGAGGTGATTAAGGTGAGGGTGGCTTCCACAGAA	8460
Qy	8461	ACCCTGAGCTGTGGGAGACAGATTTGAGCAGTGAAGAGAACCTCTGCGCTGTACAGC	8520
Db	8461	ACCCTGAGCTGTGGGAGACAGATTTGAGCAGTGAAGAGAACCTCTGCGCTGTACAGC	8520
Qy	8521	AGGACTTTGAGAGCAATGCTGCGCTGCACCTTGACATTAATATGCAATTTTCTTACCC	8580
Db	8521	AGGACTTTGAGAGCAATGCTGCGCTGCACCTTGACATTAATATGCAATTTTCTTACCC	8580
Qy	8581	TTTCTGCTCAGAAATGTCTCAGGGTCCCCCAAGGCCCTTACCA	8622
Db	8581	TTTCTGCTCAGAAATGTCTCAGGGTCCCCCAAGGCCCTTACCA	8622

Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Kelle, C., Lacombe, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKean, P., McEwan, K., Meldrum, J., Meneus, L., Mihova, T., Mlangi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trifilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 91923)
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Kelle, C., Lander, S., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlangi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 91923)
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Kelle, C., Lander, S., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlangi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 3, 2002 this sequence version replaced g1:22474954.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L24596
 Center clone name: 3038_L12

 Only the first 91.9 kb of this clone are being submitted.
 The remainder overlaps accession number AC055860 (WICGR project L9785).

FEATURES	Location/Qualifiers
source	1..91923
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11"
	/clone="CTD-3038L12"
	/clone_11b="CTD2 Human BAC"
repeat_region	25..70
	/rpt_family="L2"
	/complement(251..899)
repeat_region	/rpt_family="L2"
	900..1001
	/rpt_family="L1MD2"
repeat_region	1082..1655
	/rpt_family="L1MD2"
repeat_region	1671..1850
	/rpt_family="L1MD3"
	/complement(1858..1917)
repeat_region	1918..2273
	/rpt_family="L2"
repeat_region	/rpt_family="THE1B"
	/complement(2274..2327)
repeat_region	4505..4713
	/rpt_family="L2"
repeat_region	5536..5795
	/rpt_family="L2"
repeat_region	5796..5837
	/rpt_family="A1USC"
repeat_region	5838..5875
	/rpt_family="TAAAn"
repeat_region	/rpt_family="A1USC"
	/complement(7186..7526)
repeat_region	/rpt_family="L1MB4"
	/complement(10716..10825)
repeat_region	/rpt_family="MIR"
	11365..11466
repeat_region	/rpt_family="L2"
	/complement(11479..11644)
repeat_region	/rpt_family="MIR3"
	12072..12409
repeat_region	/rpt_family="MIR2"
	12692..12875
repeat_region	/rpt_family="L1MB1"
	/complement(12876..13294)
repeat_region	/rpt_family="MST1"
	13295..13552
repeat_region	/rpt_family="L1MB1"
	/complement(14619..14992)
repeat_region	/rpt_family="L1MB1"
	/complement(15002..17751)
repeat_region	/rpt_family="L1MB1"
	17752..19590
repeat_region	/rpt_family="L1PA6"
	19591..20184
repeat_region	/rpt_family="L1PA4"
	20185..24476
repeat_region	/rpt_family="L1PA6"
	24611..24637
repeat_region	/rpt_family="Tn"
	/complement(24639..26213)
repeat_region	/rpt_family="SVA"
	26219..26361
repeat_region	/rpt_family="L1MB5"

```
repeat_region 26372. 26560
/r/cp_family="FLAM_C"
repeat_region complement(26636. 26719)
/r/cp_family="MIR"
repeat_region complement(26663. 27139)
/r/cp_family="MER41C"
repeat_region complement(27140. 27449)
/r/cp_family="AluSg"
repeat_region complement(27450. 27494)
/r/cp_family="MER41C"
repeat_region complement(27495. 27796)
/r/cp_family="AluSg"
repeat_region complement(27797. 28030)
/r/cp_family="MER41C"
repeat_region complement(28194. 28287)
/r/cp_family="MIR"
repeat_region complement(29002. 29304)
/r/cp_family="AluDo"
repeat_region complement(29337. 29755)
/r/cp_family="MSTA"
repeat_region 30412. 30452
/r/cp_family="Tg(C) n"
repeat_region complement(31075. 31197)
/r/cp_family="MIR"
repeat_region 31198. 31246
/r/cp_family="CA" n"
repeat_region complement(31247. 31341)
/r/cp_family="MIR"
repeat_region 31541. 31821
/r/cp_family="AluSx"
repeat_region 31882. 31850
/r/cp_family="(GA) n"
repeat_region 32050. 32230
/r/cp_family="MER5A"
repeat_region complement(32650. 33747)
```

Db	15444	AGCAAAAGACAAACCCAGAGTGAAGAAAAATCTTCACAAACTAAGACTGACCAAGG	15385
Qy	1621	ACTATATCCGGAATCCACAAGAACTCAAAACAATCAGACAGAAAGAAAGCAAACTCC	1680
Db	15384	ACTAATATCCGGAATCCACAAGAACTCAAAACAATCAGACAGAAAGAAAGCAAACTCC	15325
Qy	1681	CATGAAGAAGTGGCTTAAGACATGAATAGACAATTCGCAAAAAGAAATATCAAAATGGC	1740
Db	15324	CATGAAGAAGTGGCTTAAGACATGAATAGACAATTCGCAAAAAGAAATATCAAAATGGC	15265
Qy	1741	CAACAACAGAAAAAATCTTACATCACTAATGAATGGAATGTAAATCAACACTG	1800
Db	15264	CAACAACAGAAAAAATCTTACATCACTAATGAATGGAATGTAAATCAACACTG	15205
Qy	1801	TAATGCGATACCACTTACTCTCGCAAGAATGCTCATTAATTTAAATCTAAATTAATA	1860
Db	15204	TAATGCGATACCACTTACTCTCGCAAGAATGCTCATTAATTTAAATCTAAATTAATA	15145
Qy	1861	GATGTTGGTGGTCTGTGTGTGAATAAGAAACATTTTACACTGCTGGTGGAAATGTAAC	1920
Db	15144	GATGTTGGTGGTCTGTGTGTGAATAAGAAACATTTTACACTGCTGGTGGAAATGTAAC	15085
Qy	1921	TTGGCGAACCATATGAAAAACAGTGTGAATTTCTTAAAGAACTPAAAAGTATGAC	1980
Db	15084	TTGGCGAACCATATGAAAAACAGTGTGAATTTCTTAAAGAACTPAAAAGTATGAC	15025
Qy	1981	CATTGATCCAGCAATCCCATTAATATGATTAATATATATTAATTAATACATGGAAT	2040
Db	15024	CATTGATCCAGCAATCCCATTAATATGATTAATATATATTAATTAATACATGGAAT	14965
Qy	2041	ACAACTCAGCCATAAAAAAGAAATGAATGAATGAATGAATGAATGAATGAATGAAT	2100
Db	14964	ACAACTCAGCCATAAAAAAGAAATGAATGAATGAATGAATGAATGAATGAATGAAT	14905
Qy	2101	AGACCTTATTTCTAAGTGGGGTAACTCAGAAATGGAAGAAACCAACATCATATGTTCTAC	2160
Db	14904	AGACCTTATTTCTAAGTGGGGTAACTCAGAAATGGAAGAAACCAACATCATATGTTCTAC	14845
Qy	2161	TTACAAGTGGGGCTTAAGCTGTGAGACACGAAGGCAATGAATGAATGAATGAATCTG	2220
Db	14844	TTACAAGTGGGGCTTAAGCTGTGAGACACGAAGGCAATGAATGAATGAATGAATCTG	14785
Qy	2221	GGACTTGAAGGGAAGGATGAAGAGAGCGAGGGAATAAAAGTACACATGAGGTACAGT	2280
Db	14784	GGACTTGAAGGGAAGGATGAAGAGAGCGAGGGAATAAAAGTACACATGAGGTACAGT	14725
Qy	2281	GTAACCTGCTCAGGTATGAGGTGACACCAAAATCTCAAAATTAACCACTAAAGACTTATC	2340
Db	14724	GTAACCTGCTCAGGTATGAGGTGACACCAAAATCTCAAAATTAACCACTAAAGACTTATC	14665
Qy	2341	CATGGAAGCAAAACACCACTGTTCCCAAAAATCCCATGAAATPAAAATTAATTAATAA	2400
Db	14664	CATGGAAGCAAAACACCACTGTTCCCAAAAATCCCATGAAATPAAAATTAATTAATAA	14605
Qy	2401	ATGATTTAATTTCAAGAAATTTAAAAAGTTCACTGTTCAAGATTTAATAATGAAGTA	2460
Db	14604	ATGATTTAATTTCAAGAAATTTAAAAAGTTCACTGTTCAAGATTTAATAATGAAGTA	14545
Qy	2461	AGAAATGAAGAAAGTGTAGCAAGTGTAGCTTGTGACATGAGGACTTGAATTTCACTTG	2520
Db	14544	AGAAATGAAGAAAGTGTAGCAAGTGTAGCTTGTGACATGAGGACTTGAATTTCACTTG	14485
Qy	2521	CATACACTCTCTGGCATTTGGAAGAAAGATATACATGAATATATCACCACTATGATA	2580
Db	14484	CATACACTCTCTGGCATTTGGAAGAAAGATATACATGAATATATCACCACTATGATA	14425
Qy	2581	AAAGAAACATCAAAAAATTTGTGTAGGCAATGTGTGAGCTTGAATGCTCCATGATCTAC	2640
Db	14424	AAAGAAACATCAAAAAATTTGTGTAGGCAATGTGTGAGCTTGAATGCTCCATGATCTAC	14365
Qy	2641	TTTTTCATTGGATATAAAGGCTCATATATGATGTTCACTTGTCTTAATGTATGCTTAG	2700
Db	14364	TTTTTCATTGGATATAAAGGCTCATATATGATGTTCACTTGTCTTAATGTATGCTTAG	14305
Qy	2701	GCCCATTAATTTAATTTAAATCAGGACAGCAATTAATTAATCAGAAAGTTGAACAAGATGG	2760
Db	14304	GCCCATTAATTTAATTTAATTTAATCAGGACAGCAATTAATTAATCAGAAAGTTGAACAAGATGG	14245
Qy	2761	GACGTATGAGAGAGGCTTAAATGTAATGTAATGGAATGGAACAGAGGCCAAGAAATCATCTCAG	2820
Db	14244	GACGTATGAGAGAGGCTTAAATGTAATGTAATGGAATGGAACAGAGGCCAAGAAATCATCTCAG	14185
Qy	2821	TTTGAATTTGTGTCTCAATTAACCTCTGGCTCTGATTTTGGCCATAGTCTCTCATAGGAA	2880
Db	14184	TTTGAATTTGTGTCTCAATTAACCTCTGGCTCTGATTTTGGCCATAGTCTCTCATAGGAA	14125
Qy	2881	ATPACAAAGCTGTCCAGCATCTTCCGTAAGCTGTGATGCTCACAGCTTTCATTTCACT	2940
Db	14124	ATPACAAAGCTGTCCAGCATCTTCCGTAAGCTGTGATGCTCACAGCTTTCATTTCACT	14065
Qy	2941	CCTGTAGGCACTCTCTGAATTAAGCAACAGAAAAATCTCTGAAGTCACTGAATCCCA	3000
Db	14064	CCTGTAGGCACTCTCTGAATTAAGCAACAGAAAAATCTCTGAAGTCACTGAATCCCA	14005
Qy	3001	GAAAGCTCTTACCTTTAGCAACAAGGAGCTTCAACACTGGAACAAAGAAAGCAAT	3060
Db	14004	GAAAGCTCTTACCTTTAGCAACAAGGAGCTTCAACACTGGAACAAAGAAAGCAAT	13945
Qy	3061	AAAGGTAAGTACCAAGAACTCTCTTCCAGTCAAGTCAATGATTTTGTGTAAGATC	3120
Db	13944	AAAGGTAAGTACCAAGAACTCTCTTCCAGTCAAGTCAATGATTTTGTGTAAGATC	13885
Qy	3121	ATGTCTTATGCTTCCACTTGGTGTACATGACAGAGGGGTCAAGACTTGTTCAGAAA	3180
Db	13884	ATGTCTTATGCTTCCACTTGGTGTACATGACAGAGGGGTCAAGACTTGTTCAGAAA	13825
Qy	3181	AGACAGAGACATGAAGCTTCTTCCAGAACTGAGTGTGTCAACCCAACTGTGTGAG	3240
Db	13824	AGACAGAGACATGAAGCTTCTTCCAGAACTGAGTGTGTGTCAACCCAACTGTGTGAG	13765
Qy	3241	CTCTAATGAGTGTCCCTCTTCAATTTATCTCCCATATCACTTCTTCAATTCATCA	3300
Db	13764	CTCTAATGAGTGTCCCTCTTCAATTTATCTCCCATATCACTTCTTCAATTCATCA	13705
Qy	3301	TTCAATCTGCTCATGAGAGACGTGCTCTTTACATTTAATTAAGAGCAAGGAGAC	3360
Db	13704	TTCAATCTGCTCATGAGAGACGTGCTCTTTACATTTAATTAAGAGCAAGGAGAC	13645
Qy	3361	ATGCAGGCAATTTCTTCCAGAGTGAACGTATAGAGCACTTCTTGTTCACCTAC	3420
Db	13644	ATGCAGGCAATTTCTTCCAGAGTGAACGTATAGAGCACTTCTTGTTCACCTAC	13585
Qy	3421	TTTTCAAAATTTATCTTCTTCTGCTATCTGAAAAGGCTTAAGAAATATAGTGGCCA	3480
Db	13584	TTTTCAAAATTTATCTTCTTCTGCTATCTGAAAAGGCTTAAGAAATATAGTGGCCA	13525
Qy	3481	ATATATTAAGAGTGTTCATGAGAAAGTATTTACAAAGATGACAGATTAAGGCTCAG	3540
Db	13524	ATATATTAAGAGTGTTCATGAGAAAGTATTTACAAAGATGACAGATTAAGGCTCAG	13465
Qy	3541	GATCTTAAGCAATACATATAGGGAGCACTACTTCTCCCTAGCTGAAACGAGACAGG	3600
Db	13464	GATCTTAAGCAATACATATAGGGAGCACTACTTCTCCCTAGCTGAAACGAGACAGG	13405
Qy	3601	GAAGAGCAGTTACCAATGTGCTCATATAGCTATGAGCATTAAGGTTGGAGAGCAT	3660
Db	13404	GAAGAGCAGTTACCAATGTGCTCATATAGCTATGAGCATTAAGGTTGGAGAGCAT	13345
Qy	3661	GAGCAGGCAAGTGAAGAGCCTGTGGTGCACACGACAGCCACACAGGCTGATATAGTT	3720
Db	13344	GAGCAGGCAAGTGAAGAGCCTGTGGTGCACACGACAGCCACACAGGCTGATATAGTT	13285
Qy	3721	TGGAATGTTGTTCCCAAAATCTCATGTATATTTCAATGTGAGAGAGAGG	3780
Db	13284	TGGAATGTTGTTCCCAAAATCTCATGTATATTTCAATGTGAGAGAGAGG	13225

QY	3781	CTGTGGGAGATGATTAATAGAACAGGGGATGGTTTGACATGAATGTTTAAACCAT	3840
Db	13224	CTTGTGGAGATGATTAATAGAACAGGGGATGGTTTGACATGAATGTTTAAACCAT	13165
QY	3841	CCCCCTTGGATTTGTTGTGTGATCTGACGAGTTCTCATGAAATCTAGTTTAAAA	3900
Db	13164	CCCCCTTGGATTTGTTGTGTGATCTGACGAGTTCTCATGAAATCTAGTTTAAAA	13105
QY	3901	GCGGTGACACTCTCTCCCTCTCTCTTAATCTCTGCTCAACATGTGACGCGCTGCTC	3960
Db	13104	GCGGTGACACTCTCTCCCTCTCTCTTAATCTCTGCTCAACATGTGACGCGCTGCTC	13045
QY	3961	CCCCCTTGGCTTACACAGATTGGGAAGCTTCTGAGGCGCTCCCGAAGACAAACCTGC	4020
Db	13044	CCCCCTTGGCTTACACAGATTGGGAAGCTTCTGAGGCGCTCCCGAAGACAAACCTGC	12985
QY	4021	TATGCTTCTGTAGACTGTGTAGAGACTATTAGCCAGTTAAACCATTTCTTCATAAAT	4080
Db	12984	TATGCTTCTGTAGACTGTGTAGAGACTATTAGCCAGTTAAACCATTTCTTCATAAAT	12922
QY	4081	TCCCACTCTCAGTAATTTCTTTTAAAGAAATTTGAAATGAATATACACAGACAGAG	4140
Db	12924	TCCCACTCTCAGTAATTTCTTTTAAAGAAATTTGAAATGAATATACACAGACAGAG	12865
QY	4141	CCAGAGATGGAATTCGCCAGGTGCTTTCCTGCTGTCTTCACTCTCTGCTGTCTC	4200
Db	12864	CCAGAGATGGAATTCGCCAGGTGCTTTCCTGCTGTCTTCACTCTCTGCTGTCTC	12805
QY	4201	CCAGTGTCTCAATTTCCACAGAAACAGAAATAAAAAGAAATCCCATGATGTGTACAT	4260
Db	12804	CCAGTGTCTCAATTTCCACAGAAACAGAAATAAAAAGAAATCCCATGATGTGTACAT	12744
QY	4261	GAAGCCACTCTCTTGGGATGTCAACAGATTAAGAAAGAAATGAAACAAATCTCATGG	4320
Db	12744	GAAGCCACTCTCTTGGGATGTCAACAGATTAAGAAAGAAATGAAACAAATCTCATGG	12685
QY	4321	TAAATGAGACTATCCCTCTCACTTCTGTATCTCTTAATCTGCGGGCTTCTCATTC	4380
Db	12684	TAAATGAGACTATCCCTCTCACTTCTGTATCTCTTAATCTGCGGGCTTCTCATTC	12622
QY	4381	TGATTGATCCCTGTCTCATTTAGCTGTATACAGCTATCACTTAATGGTTGGCTGTCTTC	4440
Db	12624	TGATTGATCCCTGTCTCATTTAGCTGTATACAGCTATCACTTAATGGTTGGCTGTCTTC	12565
QY	4441	TCTACTGTCACTTTATGACAGAAATGTTTGCATTTGTTAAAAATGCATAGAAATATAAT	4500
Db	12564	TCTACTGTCACTTTATGACAGAAATGTTTGCATTTGTTAAAAATGCATAGAAATATAAT	12505
QY	4501	GTAATTTTAAAAAGACATATGTAATTTTGTTAAGATATATAGTTGGCTGATCTAATAA	4560
Db	12504	GTAATTTTAAAAAGACATATGTAATTTTGTTAAGATATATAGTTGGCTGATCTAATAA	12445
QY	4561	GACATGAAAGAAATATCTTAAACAAGAAATATAGTTGGCTGTGGCTCATAGTT	4620
Db	12444	GACATGAAAGAAATATCTTAAACAAGAAATATAGTTGGCTGTGGCTCATAGTT	12385
QY	4621	CTGATCTACAGATTCACAAACCTACAGAGAAACCTTTCACAAAAATAAAGGTGGCG	4680
Db	12384	CTGATCTACAGATTCACAAACCTACAGAGAAACCTTTCACAAAAATAAAGGTGGCG	12322
QY	4681	GAGTTGTGTATGTACTGAACAGGTACAACTGTATTTCTTGTCAATTAATTTCTGAAAA	4740
Db	12324	GAGTTGTGTATGTACTGAACAGGTACAACTGTATTTCTTGTCAATTAATTTCTGAAAA	12265
QY	4741	CTAACAATTAACAAGAACTTATATAGATTTGCAATTTGTGCAGTATTTCTAATACTTT	4800
Db	12264	CTAACAATTAACAAGAACTTATATAGATTTGCAATTTGTGCAGTATTTCTAATACTTT	12205
QY	4801	AAATGATTTAATGATCTGGAGAAATGCAATAGATATATACAAATACATATATAGG	4860
Db	12204	AAATGATTTAATGATCTGGAGAAATGCAATAGATATATACAAATACATATATAGG	12145
QY	4861	AAATTGAGCATCTGCAATTTTGGTCTGTGCTGGGGTCTGAAAGAAATCCCTGTAAAT	4920

Dp	12144	AAATTGACATCTGACGATTTTGGTCTGGCTGGGGTCTTGGAAAGAAATCCCTGTAAT	12085
Qy	4921	ACACAAAATGACACTCTTGAGATCTGAACTGAAAGCTCCAAAGCATCATACACAGA	4980
Dp	12084	ACACAAAATGACACTCTTGAGATCTGAACTGAAAGCTCCAAAGCATCATACACAGA	12025
Qy	4981	TTCCAAAAATTTGCTGCTCCCGCACTTCTTAGAGAGTTGGCCCTCATCTTGTATCTCACT	5040
Dp	12024	TTCCAAAAATTTGCTGCTCCCGCACTTCTTAGAGAGTTGGCCCTCATCTTGTATCTCACT	11965
Qy	5041	GGTTCCAGGACATTAGCATTCACGCTTAATGAAAAAGAGCGAGGGGAAGAGAGGCT	5100
Dp	11964	GGTTCCAGGACATTAGCATTCACGCTTAATGAAAAAGAGCGAGGGGAAGAGAGGCT	11905
Qy	5101	TTGCTCCTTTATTTAAATCCCATGAGCCAGAGACTTGGCTTGTCTACATTTTGTATCTTCC	5160
Dp	11304	TTGCTCCTTTATTTAAATCCCATGAGCCAGAGACTTGGCTTGTCTACATTTTGTATCTTCC	11845
Qy	5161	ACTTAAACAGACCTGCTCATGGAGTGTATCCAGCATCAAGAAAACTGGAGTGGGCT	5220
Dp	11844	ACTTAAACAGACCTGCTCATGGAGTGTATCCAGCATCAAGAAAACTGGAGTGGGCT	11785
Qy	5221	CTTGTGCTGCTTTGTACATTTCTCAGAAAGGTATGTGACCAAAAAAGAAATCTTGGGGCA	5280
Dp	11784	CTTGTGCTGCTTTGTACATTTCTCAGAAAGGTATGTGACCAAAAAAGAAATCTTGGGGCA	11725
Qy	5281	ACCAGCAGTCTCTTCAGCCCTGACCTGTCTCGATTCTGTGCTCATCAAGATTTTTC	5340
Dp	11724	ACCAGCAGTCTCTTCAGCCCTGACCTGTCTCGATTCTGTGCTCATCAAGATTTTTC	11665
Qy	5341	GGAATCTTCAGAAATATATTAATGTGTGGGGCAGAGAACGAATCTGTGTCAGGA	5400
Dp	11664	GGAATCTTCAGAAATATATTAATGTGTGGGGCAGAGAACGAATCTGTGTCAGGA	11605
Qy	5401	CTCCAGGAGACAGGGGCTGGTATTTGGAAGCTGCTCTTCAATGTTGTGAACAGAAAACTCT	5460
Dp	11604	CTCCAGGAGACAGGGGCTGGTATTTGGAAGCTGCTCTTCAATGTTGTGAACAGAAAACTCT	11545
Qy	5461	TTTAATTTCTAGGCTTAGGCTTATGTTATGTATATGAGATTAATCAATAGACAGTCT	5520
Dp	11544	TTTAATTTCTAGGCTTAGGCTTATGTTATGTATATGAGATTAATCAATAGACAGTCT	11485
Qy	5521	TTAAAGAACATCATAGACATGTTAAACAAATGCTAAATTTGTGATACCAAGTGA	5580
Dp	11484	TTAAAGAACATCATAGACATGTTAAACAAATGCTAAATTTGTGATACCAAGTGA	11425
Qy	5581	AGACAGGACATGACTTACTCTCTTAACGAGATCTTGGGTTTCATAGAGAACAAACATATCA	5640
Dp	11424	AGACAGGACATGACTTACTCTCTTAACGAGATCTTGGGTTTCATAGAGAACAAACATATCA	11365
Qy	5641	TACCATTAACCATAGAGAGCAACAAAGTTAATAGTCTGTAGGTGTATACAGAGTTCTC	5700
Dp	11364	TACCATTAACCATAGAGAGCAACAAAGTTAATAGTCTGTAGGTGTATACAGAGTTCTC	11305
Qy	5701	CTTTTCTCTCCATTTCTTTTGGGCGCATCAGAGCTGTGAGAGCTTGTCTCTCTAAGAG	5760
Dp	11304	CTTTTCTCTCCATTTCTTTTGGGCGCATCAGAGCTGTGAGAGCTTGTCTCTCTAAGAG	11245
Qy	5761	AGCTCATGATGATGCACTCACTCTGTATGCTCTCTATATCTCCAGAGAGAGATGATC	5820
Dp	11244	AGCTCATGATGATGCACTCACTCTGTATGCTCTCTATATCTCCAGAGAGAGATGATC	11185
Qy	5821	TTCTTTTCCACCTGGAGAGCTCCGCCCATGTGATCTTTGGATATTCAGAGCAAAAGTGG	5880
Dp	11184	TTCTTTTCCACCTGGAGAGCTCCGCCCATGTGATCTTTGGATATTCAGAGCAAAAGTGG	11125
Qy	5881	CTCTGTATAGGCAAAAAAGAACTCTGATTTGTTCTTAAATGCGACGCACTCACCTCTA	5940
Dp	11124	CTCTGTATAGGCAAAAAAGAACTCTGATTTGTTCTTAAATGCGACGCACTCACCTCTA	11065
Qy	5941	TTTTTCCCTTAATTCATTTTGGCTTCTCAATCTGTATCTGGAGTTGTATAGTTAATTTT	6000

Db	11064	TTTTTCCCTTATTTCATTGTCCTTCTCATCTCTATCTGAGATGGATTGTTAGTTAATTTT	11005
Qy	6001	TTTTTCAGGCCCAAAATTTTGACCTGTCAACTTTGATTTTAACTTGGAAATCACTCCCTCTACT	6060
Db	11004	TTTTTTCAGGCCCAAAATTTTGACCTGTCAACTTTGATTTTAACTTGGAAATCACTCCCTCTACT	10945
Qy	6061	TTACCCCCCTCTAAACATGTATTAATTCAGACACATAGTGGGCTGGGGCTCAAAGGGCTGGTGA	6120
Db	10944	TTACCCCCCTCTAAACATGTATTAATTCAGACACATAGTGGGCTGGGGCTCAAAGGGCTGGTGA	10885
Qy	6121	AAAAATGGATCATGAGTCAGAGCCCTGTGGGCTCAATTGATCTATTAATATATAACCC	6180
Db	10884	AAAAATGGATCATGAGTCAGAGCCCTGTGGGCTCAATTGATCTATTAATATATAACCC	10825
Qy	6181	CCCGGACAAATTAATATCTCTCTTTAACTTAATTTCACTTAATCTGGCAATACAGAAATTA	6240
Db	10824	CCCGGACAAATTAATATCTCTCTTTAACTTAATTTCACTTAATCTGGCAATACAGAAATTA	10765
Qy	6241	TACTAATTTTTTAACTCCTTAGAGCTCTTAGATGATTTAAAGAGGCAATACCTAAATAACTG	6300
Db	10764	TACTAATTTTTTAACTCCTTAGAGCTCTTAGATGATTTAAAGAGGCAATACCTAAATAACTG	10705
Qy	6301	TCAATCAGCTGTCTGTATTCTCTCCAAATTAGACCTAATCTCATCTCCAGTTGAAATTT	6360
Db	10704	TCAATCAGCTGTCTGTATTCTCTCCAAATTAGACCTAATCTCATCTCCAGTTGAAATTT	10645
Qy	6361	GCATGAATATCTCTCTTTAAACCCCAAGCCTTAACTTCTCCTATTTTCACTCATGAGACT	6420
Db	10644	GCATGAATATCTCTCTTTAAACCCCAAGCCTTAACTTCTCCTATTTTCACTCATGAGACT	10585
Qy	6421	CCCTCATATCAAAATGTTTGCATCAACAAAGAAAGGCTACCAAGATCTCCGGAAGAGAG	6480
Db	10584	CCCTCATATCAAAATGTTTGCATCAACAAAGAAAGGCTACCAAGATCTCCGGAAGAGAG	10525
Qy	6481	AATGAATATAGTGTACATTTGTATGTACTCAGACAGAAACCTTAGTAGTCCCCCATATAT	6540
Db	10524	AATGAATATAGTGTACATTTGTATGTACTCAGACAGAAACCTTAGTAGTCCCCCATATAT	10465
Qy	6541	TCCACACTTCAATTACCTGTGTGAGTGGCACTCAGGCTCAACCTTCATCTACTTTTCT	6600
Db	10464	TCCACACTTCAATTACCTGTGTGAGTGGCACTCAGGCTCAACCTTCATCTACTTTTCT	10405
Qy	6601	CTGTTCTATTGCTGAGGAATTCAGACTCAGACCCACACCTTACCMAACATGTGTACAA	6660
Db	10404	CTGTTCTATTGCTGAGGAATTCAGACTCAGACCCACACCTTACCMAACATGTGTACAA	10345
Qy	6661	ATGCTTCTAAGGGGTTTGGCAAAAGCACACTGAGTCTTATTTTAAAGCACATCAGTGT	6720
Db	10344	ATGCTTCTAAGGGGTTTGGCAAAAGCACACTGAGTCTTATTTTAAAGCACATCAGTGT	10285
Qy	6721	CAATTTCAAGTTTGGGCACTCATCATCATTTCTTCAACACAGATAGAGCTGTCCACA	6780
Db	10284	CAATTTCAAGTTTGGGCACTCATCATCATTTCTTCAACACAGATAGAGCTGTCCACA	10225
Qy	6781	AATAGAAATCTGAATGAATGAATTTTCTTCATCTAATTAATAGTGGTCTTAATGCT	6840
Db	10224	AATAGAAATCTGAATGAATGAATTTTCTTCATCTAATTAATAGTGGTCTTAATGCT	10165
Qy	6841	TACATTTGTCTTCAATTTTATTTTATTTCCATTTTCATCCAAATCTACATTTAGGCTT	6900
Db	10164	TACATTTGTCTTCAATTTTATTTTATTTCCATTTTCATCCAAATCTACATTTAGGCTT	10105
Qy	6901	CTCATCATGATCATTTCTTCATTTGAATGAACGTTTATGAAGACAAATTTGTGCTTATG	6960
Db	10104	CTCATCATGATCATTTCTTCATTTGAATGAACGTTTATGAAGACAAATTTGTGCTTATG	10045
Qy	6961	GAATAGGCACTAGAGATATAAATGTAAATGTGGTCTGTGTGCAATGACTGACACT	7020
Db	10044	GAATAGGCACTAGAGATATAAATGTAAATGTGGTCTGTGTGCAATGACTGACACT	9985
Qy	7021	GAGTTATTTTCAACCCCAAGAGTCCCGGCACTTTTCAACACATCTGTGCGAAGATCCCATTT	7080
Db	9984	GAGTTATTTTCAACCCCAAGAGTCCCGGCACTTTTCAACACATCTGTGCGAAGATCCCATTT	9925

QY	7081	TTCCCTGGGTTCAATPATSCAGATCTTTTTCCTGCTCAGAGATGACACAGCTCTGATCATG	7140
Db	9924	TTCTCTGGTTCAATPATGATGATCTTTTTCCTGCTCAGAGATGACACAGCTCTGATCATG	9865
QY	7141	AGGGTGTCAACACCCTCTTTGTGTATCTGAAATTCCTCCACTGTAGAGAAAATTTTCAGG	7200
Db	9864	AGGGTGTCAACACCCTCTTTGTGTATCTGAAATTCCTCCACTGTAGAGAAAATTTTCAGG	9805
QY	7201	CCCAAGATPATAGTATATCATGTGGGGTCCACACATCGGCTATGATAGATGAGGGGGTGTGTGAT	7260
Db	9804	CCCAAGATPATAGTATATCATGTGGGGTCCACACATCGGCTATGATAGATGAGGGGGTGTGTGAT	9745
QY	7261	CTTATATGTATATCCCATATGTACAGACAGAACTTGTGTGGCAGTATAGAGAGAGTCAAGCTTTC	7320
Db	9744	CTTATATGTATATCCCATATGTACAGACAGAACTTGTGTGGCAGTATAGAGAGAGTCAAGCTTTC	9685
QY	7321	AGAGTCAACAAGAACTGTGAATTTCAAACTGTGATTTGAGAGACCCCCACTTTTGATATAGTGA	7380
Db	9684	AGAGTCAACAAGAACTGTGAATTTCAAACTGTGATTTGAGAGACCCCCACTTTTGATATAGTGA	9625
QY	7381	CTTATATCTCTGCGAGTCTCTGATCTCTCCCTTTTAAATGAGACAGATTAATTCACATG	7440
Db	9624	CTTATATCTCTGCGAGTCTCTGATCTCTCCCTTTTAAATGAGACAGATTAATTCACATG	9565
QY	7441	CAGGGTGTGGGGAGAAATCAGAGATCAACAGACTGTGATCACATCTGTTTCTGTTTCC	7500
Db	9564	CAGGGTGTGGGGAGAAATCAGAGATCAACAGACTGTGATCACATCTGTTTCTGTTTCC	9505
QY	7501	AGGGTCAATCAAGCTGGGGGTTTCTGAGCATGTGAATTCAAACATCCCAAGCTTGGGGTACAGAA	7560
Db	9504	AGGGTCAATCAAGCTGGGGGTTTCTGAGCATGTGAATTCAAACATCCCAAGCTTGGGGTACAGAA	9445
QY	7561	CTGACACCAATCAACCGACGTGAGAGACTCTTGTCTAACAGACAGACCTGACCTTTCAG	7620
Db	9444	CTGACACCAATCAACCGACGTGAGAGACTCTTGTCTAACAGACAGACCTGACCTTTCAG	9385
QY	7621	GGGCTGACATGTGCATCTGTTTCCCTTGTGTGCGCTGACAGAAAACCGGGTGTGTGCTCTGGCTC	7680
Db	9384	GGGCTGACATGTGCATCTGTTTCCCTTGTGTGCGCTGACAGAAAACCGGGTGTGTGCTCTGGCTC	9325
QY	7681	CTGGGCTGTGCGCATGACGCCACAGGAACGCTGTCTCATCTACATCTCTCAACTGTGTGCGGGCC	7740
Db	9324	CTGGGCTGTGCGCATGACGCCACAGGAACGCTGTCTCATCTACATCTCTCAACTGTGTGCGGGCC	9265
QY	7741	GACTTTCCTCTTCTTTAGGGGCCACATTAATATGTTGCCGTTTAGGCGCTCATCATATATCCGC	7800
Db	9264	GACTTTCCTCTTCTTTAGGGGCCACATTAATATGTTGCCGTTTAGGCGCTCATCATATATCCGC	9205
QY	7801	CATCCCAATCTCCAAATCTCTCAAGCTCGTGTATGACCTTTCCCTACTTTATATAGGCGTAAAGC	7860
Db	9204	CATCCCAATCTCCAAATCTCTCAAGCTCGTGTATGACCTTTCCCTACTTTATATAGGCGTAAAGC	9145
QY	7861	ATGTGTAGCGGCATCAGACCCGAGCGGTGCTGTGCATCTGTGTGGCCCATCTGTGTACAC	7920
Db	9144	ATGTGTAGCGGCATCAGACCCGAGCGGTGCTGTGCATCTGTGTGGCCCATCTGTGTACAC	9085
QY	7921	TGCGCGCGCGCCCAATATCTGTCAATCGGTATATGTGTCTCCGTCTGTGGGGCTGTGCTCCGTG	7980
Db	9084	TGCGCGCGCGCCCAATATCTGTCAATCGGTATATGTGTCTCCGTCTGTGGGGCTGTGCTCCGTG	9025
QY	7981	CTGGCGAGATATCTGTAGGTGATATTTCTGTGACCTTCTGTGTTTAGTGTGCTGATATTCGTAT	8040
Db	9024	CTGGCGAGATATCTGTAGGTGATATTTCTGTGACCTTCTGTGTTTAGTGTGCTGATATTCGTAT	8965
QY	8041	TGTGTGTAAACGTACAGATTTCAATTACATATGCGGTGCTGGTGTTTTATATGTGTGTTCTC	8100
Db	8964	TGTGTGTAAACGTACAGATTTCAATTACATATGCGGTGCTGGTGTTTTATATGTGTGTTCTC	8905
QY	8101	TGTGTGTCAACCTGTGTCTGTGCTGTGATAGATTTCTGTGTGATATCCCGAAGATGCTCGCTG	8160
Db	8904	TGTGTGTCAACCTGTGTCTGTGCTGTGATAGATTTCTGTGTGATATCCCGAAGATGCTCGCTG	8845

QY 8161 ACCAGGCTGACGACCATCTCTCAACAGTGTGCTTCTCTCTCTGCGCTGCC 8220
 DB 8844 ACCAGGCTGACGACCATCTCTCAACAGTGTGCTTCTCTCTCTGCGCTGCC 8785
 QY 8221 TTGGGCAATGAGTGGGCGCTGTTTCCAGATCCACCGATGGAAAGCTTATTTTGT 8280
 DB 8784 TTGGGCAATGAGTGGGCGCTGTTTCCAGATCCACCGATGGAAAGCTTATTTTGT 8725
 QY 8281 CATGTGATCTAGTGTTCATTTTCTGCTGCTTAAACAGAGTGGCCACCATCATTT 8340
 DB 8724 CATGTGATCTAGTGTTCATTTTCTGCTGCTTAAACAGAGTGGCCACCATCATTT 8665
 QY 8341 TACTTCTTCTGCGCTCTTAAAGCAGTGTCAAAATAGGCAACCTGAACTGTTCTC 8400
 DB 8664 TACTTCTTCTGCGCTCTTAAAGCAGTGTCAAAATAGGCAACCTGAACTGTTCTC 8605
 QY 8401 CAGAGGCTCTGACGACACCGCTGAGGTGATGAGGTGAGGCTTCTCAGGAA 8460
 DB 8604 CAGAGGCTCTGACGACACCGCTGAGGTGATGAGGTGAGGCTTCTCAGGAA 8545
 QY 8461 ACCCTGAGCTGTGCGGAGAGATTTGAGAGCAGTGAAGACCTTGCCTGTCAAGC 8520
 DB 8544 ACCCTGAGCTGTGCGGAGAGATTTGAGAGCAGTGAAGACCTTGCCTGTCAAGC 8485
 QY 8521 AGGACTTTGAGAGCAATGCTGCTGCGACACCTTGAACATTAATGATTTTCTTACC 8580
 DB 8484 AGGACTTTGAGAGCAATGCTGCTGCGACACCTTGAACATTAATGATTTTCTTACC 8425
 QY 8581 TTTCGCTCAGAAATGCTCAGGGTCC 8609
 DB 8424 TTTCGCTCAGAAATGCTCAGGGTCC 8396

RESULT 3
 AC090099/c 172939 bp DNA linear PRI 01-OCT-2002
 LOCUS Homo sapiens chromosome 11, clone RP11-113D6, complete sequence.
 AC090099
 AC090099.14 GI:23396332
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCES
 1 (bases 1 to 172939)
 Birren, B., Nussbaum, C. and Lander, E.
 Homo sapiens chromosome 11, clone RP11-113D6
 Unpublished
 2 (bases 1 to 172939)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
 Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Hesford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Karatas, A., Lacroque, K., Lamazares, R., Landers, T.,
 Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
 Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
 McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
 Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S., Severy, P.,
 Sougnaz, C., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
 AUTHORS
 3 (bases 1 to 172939)
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gind, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuppach, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 172939)

REFERENCE
 AUTHORS
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gind, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuppach, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 1, 2002 this sequence version replaced gi:22726159.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12764
 Center clone name: L13_D_6

 Location/Qualifiers
 1. 172939
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-113D6"
 /clone_1lb="RP11-113D6 Human Male BAC"
 complement(1..77)
 /rpt_family="Alus"
 complement(154..680)
 /rpt_family="Char11e1"
 complement(678..810)
 /rpt_family="Char11e1"
 complement(879..1057)
 repeat_region
 repeat_region
 repeat_region
 repeat_region
 repeat_region

```

repeat_region /rpt_family="L1M4"
complement(1395..1513)
/rpt_family="FLAM_A"
unsure
1870..1876
/note="<30 qual SNGL region"
unsure
1902..1910
/note="<30 qual SNGL region"
repeat_region 2672..2788
/rpt_family="5S"
complement(4009..4170)
repeat_region /rpt_family="MIR"
complement(4210..4355)
repeat_region /rpt_family="MERSB"
4356..4647
/rpt_family="AluSx"
complement(6554..6658)
repeat_region /rpt_family="LTR41"
6916..6985
/rpt_family="MLT1C"
complement(7175..7359)
repeat_region /rpt_family="LTR41"
complement(8308..8611)
repeat_region /rpt_family="Aluub"
8622..9083
/rpt_family="MLT2B1"
complement(9265..10668)
repeat_region /rpt_family="L1PB3"
10699..10809
/rpt_family="MERS4"
repeat_region /rpt_family="L1M2A"
10782..11264
/rpt_family="LTR48B"
complement(11344..11683)
repeat_region /rpt_family="L2"
11835..12026
/rpt_family="L1MC4a"
12188..12660
/rpt_family="L1MA2"
repeat_region 12879..13232
/rpt_family="L1MC4a"
13223..13522
/rpt_family="AluSx"
repeat_region 13523..13817
/rpt_family="L1MC4a"
14819..15108
/rpt_family="AluJo"
repeat_region 15114..15175
/rpt_family="L1"
16150..16336
/rpt_family="MERSA"
complement(16587..16598)
/note="<30 qual SNGL region"
unsure
complement(16607..16612)
/note="<30 qual SNGL region"
repeat_region complement(16746..17360)
/rpt_family="L1ME3"
17361..17664
/rpt_family="AluJo"
repeat_region complement(17665..17774)
/rpt_family="L1ME3"
repeat_region complement(17827..17929)
/rpt_family="MIR"
repeat_region 18993..19298
/rpt_family="AluJo"
repeat_region 19755..19821
/rpt_family="L2"
19933..19941
/note="<30 qual SNGL region"
repeat_region 20320..20609
/rpt_family="AluJo"
repeat_region complement(20684..20994)
/rpt_family="AluSx"
20713..20721
/note="<30 qual SNGL region"

```

```

unsure 20879..20883
/note="<30 qual SNGL region"
repeat_region complement(21118..21398)
/rpt_family="AluSx"
unsure 21141..21147
/note="<30 qual SNGL region"
unsure 21144
/note="probably A; possibly G"
unsure 21146
/note="probably A; possibly G"
unsure 21153..21158
/note="probably A; possibly G"
unsure 21154
/note="<30 qual SNGL region"
unsure 21154
/note="probably T; possibly C"
unsure 21160..21167
/note="<30 qual SNGL region"

Query Match 99.3%; Score 8564.2; DB 9; Length 172939;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 8594; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1 TGTATGAGCCATGTCACTTTAATACCAAAACCGAAAGATTTACAAAAAGAAAC 60
DB 121497 TGTATGAGCCATGTCACTTTAATACCAAAACCGAAAGATTTACAAAAAGAAAC 121438

QY 61 TATGACAGTACCACTGATGATATATACATGAGAAATCCCAAAATACCTAGTAAC 120
DB 121437 TATGACAGTACCACTGATGATATATACATGAGAAATCCCAAAATACCTAGTAAC 121378

QY 121 CCATCCAAACAGCATATCAAGAAATATCCACATTTGCAATGGGTTTCATACAGGG 180
DB 121377 CCATCCAAACAGCATATCAAGAAATATCCACATTTGCAATGGGTTTCATACAGGG 121318

QY 181 GTGCGAGATAGTTAATCATACCAAGTCAATTAATGTGATACATCAATAAACAGATTA 240
DB 121317 GTGCGAGATAGTTAATCATACCAAGTCAATTAATGTGATACATCAATAAACAGATTA 121258

QY 241 AAAACAAAATCAACATGATCATCTCAATATGATGCTGAAAAAGCATTTGACAAAATCTAAC 300
DB 121257 AAAACAAAATCAACATGATCATCTCAATATGATGCTGAAAAAGCATTTGACAAAATCTAAC 121198

QY 301 ATTTCTTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 360
DB 121197 ATTTCTTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 121138

QY 361 TAAAAGCCATATATGACGAGCCACAGCAAAATATATATGATGAGGAAAAAGTTGAAA 420
DB 121137 TAAAAGCCATATATGACGAGCCACAGCAAAATATATATGATGAGGAAAAAGTTGAAA 121078

QY 421 CATTCCTCGAAGACTGGAACAAGAGATGCTACTTCAACCACTTCTATTCACAT 480
DB 121077 CATTCCTCGAAGACTGGAACAAGAGATGCTACTTCAACCACTTCTATTCACAT 121018

QY 481 AGTAGTGAAGTTTATGACGAGCAATCAAGCAAGAAAGAAATCAAGGGCAACCAAT 540
DB 121017 AGTAGTGAAGTTTATGACGAGCAATCAAGCAAGAAAGAAATCAAGGGCAACCAAT 120958

QY 541 CATATAAGAGAGTAAATCTGCTCTGTCCTGATATATGATTTGATTTCTTGAAGA 600
DB 120957 CATATAAGAGAGTAAATCTGCTCTGTCCTGATATATGATTTGATTTCTTGAAGA 120898

QY 601 CCTTAAGACTCATCCAGAAAGCTCCTAGAACTGATACATAAATCACTAAAGTTTCAG 660
DB 120897 CCTTAAGACTCATCCAGAAAGCTCCTAGAACTGATACATAAATCACTAAAGTTTCAG 120838

QY 661 ATACAACTAAATGTACCAAAATCACTGATACCAACCAAGTGAAGTGA 720
DB 120837 ATACAACTAAATGTACCAAAATCACTGATACCAACCAAGTGAAGTGA 120778

QY 721 GAATCAATCAAGAACTCAACACTTTTCAATAGCTGTAAAAAATACCTTAAGATAT 780
DB 120777 GAATCAATCAAGAACTCAACACTTTTCAATAGCTGTAAAAAATACCTTAAGATAT 120718

```


Qy	781	CTTACCCAGAGGTGAAGACCTCTACAGGAAACTACAAACACAGCTGACATATA	840
Db	120117	CTTACCCAGAGGTGAAGACCTCTACAGGAAACTACAAACACAGCTGACATATA	120658
Qy	841	GATGACACAAACAGTGGAAACATCCATGCTCATGATGGGTAGATCATATATTGTG	900
Db	120657	GATGACACAAACAGTGGAAACATCCATGCTCATGATGGGTAGATCATATATTGTG	120598
Qy	901	AAATGACCATATTGGCCAAAGCAATCTACAGATTCAATGCAATTTCCACCAAAATATCA	960
Db	120597	AAATGACCATATTGGCCAAAGCAATCTACAGATTCAATGCAATTTCCACCAAAATATCA	120538
Qy	961	TCATCATCTCTCCAGAACTAGGAAAAACAATCTTAAATTCCTATAGGAAACAACCA	1020
Db	120537	TCATCATCTCTCCAGAACTAGGAAAAACAATCTTAAATTCCTATAGGAAACAACCA	120479
Qy	1021	AAAAAAAAAAAAAAAACCCGCATAGCCAAAGCAAGACTTAGCAAAAAGAACTATGGA	1080
Db	120478	AAAAAAAAAAAAAAAACCCGCATAGCCAAAGCAAGACTTAGCAAAAAGAACTATGGA	120419
Qy	1081	GGCATCATTTACCATCTTCAAACTATACTACAGGCTATATCACCMAAACATCATGG	1140
Db	120418	GGCATCATTTACCATCTTCAAACTATACTACAGGCTATATCACCMAAACATCATGG	120359
Qy	1141	CATGACATATAAACTAGGCACTAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAG	1200
Db	120358	CATGACATATAAACTAGGCACTAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAG	120299
Qy	1201	CCAAATTAATATAGCCCAATGATTTTGAACAAGCAACAACAAATCTAAAGTGGGAAA	1260
Db	120298	CCAAATTAATATAGCCCAATGATTTTGAACAAGCAACAACAAATCTAAAGTGGGAAA	120239
Qy	1261	GACATCTAGTTAACAAATGCTGTGAGATTATTTGCAAGCCCATGTGGAAGATGAAA	1320
Db	120238	GACATCTAGTTAACAAATGCTGTGAGATTATTTGCAAGCCCATGTGGAAGATGAAA	120179
Qy	1321	CTGATCCCTTGTCTCTCACTTAATACAAAATTGATACAAAGATGATCAAAAGCTTAA	1380
Db	120178	CTGATCCCTTGTCTCTCACTTAATACAAAATTGATACAAAGATGATCAAAAGCTTAA	120120
Qy	1381	TCAGAGACTTAAACCATTAATAAAATTTAGAAATACATCAGAAAATGCTTACAGCAT	1440
Db	120119	TCAGAGACTTAAACCATTAATAAAATTTAGAAATACATCAGAAAATGCTTACAGCAT	120060
Qy	1441	TCATTAAGCAAGACTTCAATGCGCAAGAACCCMAAGTAAATGCAACAAAACAAAAT	1500
Db	120059	TCATTAAGCAAGACTTCAATGCGCAAGAACCCMAAGTAAATGCAACAAAACAAAAT	120000
Qy	1501	AAATAGATAGACTTAATTTAACTTAAAGCTTTTGGCGACAAAACATCATTTAGCAG	1560
Db	119999	AAATAGATAGACTTAATTTAACTTAAAGCTTTTGGCGACAAAACATCATTTAGCAG	119940
Qy	1561	AGCAAAACAGAACCCACCGAGTGAAGAAATCTTCAAACTTAGCATCTGACTAAGG	1620
Db	119993	AGCAAAACAGAACCCACCGAGTGAAGAAATCTTCAAACTTAGCATCTGACTAAGG	119880
Qy	1621	ACTAATATTCGGAATCCACAAGAACTCAAACAATCAGCAAGAAAGAAAGCAACATCC	1680
Db	119879	ACTAATATTCGGAATCCACAAGAACTCAAACAATCAGCAAGAAAGAAAGCAACATCC	119820
Qy	1681	CATGAAAGAGTGGCTTAAGACATGAAATAGCAATTTCTCAAAAGAAATATCAAAATGGC	1740
Db	119819	CATGAAAGAGTGGCTTAAGACATGAAATAGCAATTTCTCAAAAGAAATATCAAAATGGC	119760
Qy	1741	CAACAAACAGAAAAAATGCTTAACTCATATGATTTAGGAAATGTAATCAACCTG	1800
Db	119759	CAACAAACAGAAAAAATGCTTAACTCATATGATTTAGGAAATGTAATCAACCTG	119700
Qy	1801	TAAATGCAATCAACCTTACTCTGCAAGAAATGTCATATTTTAAATCTTAAATATATA	1860
Db	119699	TAAATGCAATCAACCTTACTCTGCAAGAAATGTCATATTTTAAATCTTAAATATATA	119640
Qy	1861	GATGTTGGTGGTCTGTGTGTATAAAGAACATTTTACACTGTGTGGGAATGTAAAC	1920

Db	119639	GATGTTGGTGGTCTGTGTGTATAAAGAACATTTTACACTGTGTGGGAATGTAAAC	119580
Qy	1921	TTGGCAACCACTATGAAAAACAGTGTGAATTTCTTAAAGAACTTAAAGTATGACAC	1980
Db	119579	TTGGCAACCACTATGAAAAACAGTGTGAATTTCTTAAAGAACTTAAAGTATGACAC	119520
Qy	1981	CATTGATCCAGCAATCCCATTAATATGTATTAATATATATATTTATATCCATGCAAT	2040
Db	119519	CATTGATCCAGCAATCCCATTAATATGTATTAATATATATATTTATATCCATGCAAT	119460
Qy	2041	ACAACCTGAGCATTAATAAAGAAATTAATGATGACATTCACAGCAATCTAGTGAATTTG	2100
Db	119459	ACAACCTGAGCATTAATAAAGAAATTAATGATGACATTCACAGCAATCTAGTGAATTTG	119400
Qy	2101	AGACCTTATTTCTAAGTGGGTAACCTGAGAAATGGAACCAACATCATATGTTCTCAC	2160
Db	119399	AGACCTTATTTCTAAGTGGGTAACCTGAGAAATGGAACCAACATCATATGTTCTCAC	119340
Qy	2161	TTACAAAGTGGGCTAAGCTGTGAGACACGAAAGCATTAAGATATATGAATCTCTGG	2220
Db	119339	TTACAAAGTGGGCTAAGCTGTGAGACACGAAAGCATTAAGATATATGAATCTCTGG	119280
Qy	2221	GGACTTGAAGGGAAGGATGGAAGAGAGGCAAGGATTAAGACTACCAATGGGTACAGT	2280
Db	119279	GGACTTGAAGGGAAGGATGGAAGAGAGGCAAGGATTAAGACTACCAATGGGTACAGT	119220
Qy	2281	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTTACCTAATAAGACTTATC	2340
Db	119219	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTTACCTAATAAGACTTATC	119160
Qy	2341	CATGGAAGCAACCACTGTTCCCAAAATCCCAATGAATTAATAATAATAATAATA	2400
Db	119159	CATGGAAGCAACCACTGTTCCCAAAATCCCAATGAATTAATAATAATAATAATAATA	119100
Qy	2401	ATGATTTAATTTTCAACAATTTTAAATAAGTTCCTGTTGAGATTATTAATAAGAAATA	2460
Db	119099	ATGATTTAATTTTCAACAATTTTAAATAAGTTCCTGTTGAGATTATTAATAAGAAATA	119040
Qy	2461	AGATGAAAAAGTATGAGCAATGTAGCTCTGAGCAATGGAACCTAGATTTTACCTTG	2520
Db	119039	AGATGAAAAAGTATGAGCAATGTAGCTCTGAGCAATGGAACCTAGATTTTACCTTG	118980
Qy	2521	CATACACTTCTGCGCAATTTGAAAGAAAGTATACATGAATATATACCATATATATA	2580
Db	118979	CATACACTTCTGCGCAATTTGAAAGAAAGTATACATGAATATATACCATATATATA	118920
Qy	2581	AAAGAAACATCAAAAAATTTGTGCAGGCCATTTGCAGCCTTGAATGCTCCATGATCAC	2640
Db	118919	AAAGAAACATCAAAAAATTTGTGCAGGCCATTTGCAGCCTTGAATGCTCCATGATCAC	118860
Qy	2641	TTTTTCAATTTGGAATATAAGCCTCATATATAGTATGATGATGATGATGATGATGATG	2700
Db	118859	TTTTTCAATTTGGAATATAAGCCTCATATATAGTATGATGATGATGATGATGATGATG	118800
Qy	2701	GCCCATTAATGATTTTAAATACAGACAGCAATTAATCTTACAGAACTTGAACAAATGG	2760
Db	118799	GCCCATTAATGATTTTAAATACAGACAGCAATTAATCTTACAGAACTTGAACAAATGG	118740
Qy	2761	GAGGTATAGAGAGAGGCTTAAATATGTACTGATATGGAACAGAGGCCAAGAAATCATCTCAG	2820
Db	118739	GAGGTATAGAGAGAGGCTTAAATATGTACTGATATGGAACAGAGGCCAAGAAATCATCTCAG	118680
Qy	2821	TTAGATTTGTCTCAATATACCTGCGCTCTGATTTGGCCATAGTCTCATACAGAAA	2880
Db	118679	TTAGATTTGTCTCAATATACCTGCGCTCTGATTTGGCCATAGTCTCATACAGAAA	118620
Qy	2881	ATAACAAAGCTGTCCAGCATCTTGTATAGCTTGATGCTTCAACAGCTTTTCAATTTCACT	2940
Db	118619	ATAACAAAGCTGTCCAGCATCTTGTATAGCTTGATGCTTCAACAGCTTTTCAATTTCACT	118560
Qy	2941	CTGTAGGCACTCTCGAATTAAGCAACAGAAAAGTCTCTGAAATCTCACTGAATCCCA	3000

Db 118559 CCTGATGAGATCTCTGATTAAGCAACAGAAAAGTCTCTGATGACTGATATCCCA 118500
Qy 3001 GAAAGGCTCTTACTTATAGCAACAGGAGGCTTCAACACTGAGCAAAAGAGAAAGAT 3060
Db 118499 TAAAGGCTCTTACTTATAGCAACAGGAGGCTTCAACACTGAGCAAAAGAGAAAGAT 118440
Qy 3061 AAGGTAAGTACCAAGAACTCTCTTCAACAGTGTATGATTTTGTGCTAAGATC 3120
Db 118439 AAGGTAAGTACCAAGAACTCTCTTCAACAGTGTATGATTTTGTGCTAAGATC 118380
Qy 3121 ATGCTTATGCTTCAACCTGTGTCTACATGACAGGAGGCTCAAGACTTGTTCAGAAA 3180
Db 118379 ATGCTTATGCTTCAACCTGTGTGTCTACATGACAGGAGGCTCAAGACTTGTTCAGAAA 118320
Qy 3181 AGACAGGAGACATGAAAGCTTCTTCAAAAAGTGTGTCTCAACCCAACTGTGTGAG 3240
Db 118319 AGACAGGAGACATGAAAGCTTCTTCAAAAAGTGTGTCTCAACCCAACTGTGTGAG 118260
Qy 3241 CTCTAAATGAGTGTCCCTCTTCAATTTATCTCCCAATATCACTGCTTCAATCAATCA 3300
Db 118259 CTCTAAATGAGTGTCCCTCTTCAATTTATCTCCCAATATCACTGCTTCAATCAATCA 118200
Qy 3301 TTCAATCTGCTTCAAG 3360
Db 118199 TTCAATCTGCTTCAAG 118140
Qy 3361 ATGCAAGCATTTCTTCCAGAGTGTGAAGCTATAGAGCAAGTTCTTGTCTTCACTTAC 3420
Db 118139 ATGCAAGCATTTCTTCCAGAGTGTGAAGCTATAGAGCAAGTTCTTGTCTTCACTTAC 118080
Qy 3421 TTTTCAATTTATCTTCTTGTGCTATCTGAAAAGGCTTAAGAGATATAGATGAGCCCA 3480
Db 118079 TTTTCAATTTATCTTCTTGTGCTATCTGAAAAGGCTTAAGAGATATAGATGAGCCCA 118020
Qy 3481 ATAATTAAGAGAGTGTTCATGAGAAAGATTTTCAAAAGTGCACAGAGTTAAGGGTCAAG 3540
Db 118019 ATAATTAAGAGAGTGTTCATGAGAAAGATTTTCAAAAGTGCACAGAGTTAAGGGTCAAG 117960
Qy 3541 GATCTTAAGAGAGATATCATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 117959 GATCTTAAGAGAGATATCATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 117900
Qy 3601 GAAAGAGAGAGTACCATGTGTGCGCATAGCATAGTGTGAGCCATAAGAGGAGAGAGAT 3660
Db 117899 GAAAGAGAGAGTACCATGTGTGCGCATAGCATAGTGTGAGCCATAAGAGGAGAGAT 117840
Qy 3661 GAGCAGGAGAGTGAAGAAAGCCCTGCGTGGCAACGACAGCCACAGAGCTGATATAGTT 3720
Db 117839 GAGCAGGAGAGTGAAGAAAGCCCTGCGTGGCAACGACAGCCACAGAGCTGATATAGTT 117780
Qy 3721 TGGATCTGTGTCTCCCAACCAAAATCTCATGTGTGATTTTCAATGTGTGAGAAAGGG 3780
Db 117779 TGGATCTGTGTCTCCCAACCAAAATCTCATGTGTGATTTTCAATGTGTGAGAAAGGG 117720
Qy 3781 CCTTGGAGAGATGATTTATGATGACGAGGAGAGGTTTGGATGAAATGTTTAAACCAT 3840
Db 117719 CCTTGGAGAGATGATTTATGATGACGAGGAGAGGTTTGGATGAAATGTTTAAACCAT 117660
Qy 3841 CCCCTTGGATGATGTGTGTGTGATGATGACGAGTTCTCATGAAATCTAGTTGTTTAAAA 3900
Db 117659 CCCCTTGGATGATGTGTGTGTGATGATGACGAGTTCTCATGAAATCTAGTTGTTTAAAA 117600
Qy 3901 GGGTGTAGAGACTCTCTCTCTCTCTTACTCTGCTGTCTCAACATGTGAGAGCGCTGCTC 3960
Db 117599 GGGTGTAGAGACTCTCTCTCTCTCTTACTCTGCTGTCTCAACATGTGAGAGCGCTGCTC 117540
Qy 3961 CCCCTTGGCTTTTCAACAGAGATGAAAGCTTCTGAGGCTCCCAAGAGAGAGAGTGC 4020
Db 117539 CCCCTTGGCTTTTCAACAGAGATGAAAGCTTCTGAGGCTCCCAAGAGAGAGAGTGC 117480
Qy 4021 TATGCTTCTTGTATACAGTGTGTATAGAGTATTAAGCAGTTAAACCATTTCTTATTAAT 4080
Db 117479 TATGCTTCTTGTATACAGTGTGTATAGAGTATTAAGCAGTTAAACCATTTCTTATTAAT 117420

Qy 4081 TCCCACTCAGATATTTCTTTTATGCAATTTGAAATGAATCTATATACAGAGAGAG 4140
Db 117419 TCCCACTCAGATATTTCTTTTATGCAATTTGAAATGAATCTATATACAGAGAGAG 117360
Qy 4141 CCAGAGATGAGAAATCCCAAGGAGGCTTCTGCTGTCTCTCCAGTCTCTGCTGTGTCTC 4200
Db 117359 CCAGAGATGAGAAATCCCAAGGAGGCTTCTGCTGTCTCTCCAGTCTCTGCTGTGTCTC 117300
Qy 4201 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCATGATGTGTACTTA 4260
Db 117299 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCATGATGTGTACTTA 117240
Qy 4261 GAAAGCAGCTCTGTGGAGAGTCAAAACAGATTAAGAAAGTGAAGAAATCCCATAGG 4320
Db 117239 GAAAGCAGCTCTGTGGAGAGTCAAAACAGATTAAGAAAGTGAAGAAATCCCATAGG 117180
Qy 4321 TAAATGAGACTATCCCTCTCACCTTCTGTATCTCTTAATTCCTGAGGCTTCTCTATC 4380
Db 117179 TAAATGAGACTATCCCTCTCACCTTCTGTATCTCTTAATTCCTGAGGCTTCTCTATC 117120
Qy 4381 TGAATGATCCCTGTCTCATTTCACTCTATCACTACTTAAATGTTGGCTGTCTTTC 4440
Db 117119 TGAATGATCCCTGTCTCATTTCACTCTATCACTACTTAAATGTTGGCTGTCTTTC 117060
Qy 4441 TCTACTGTCACTTTATGAGAAATGTTGCAATTTGTTAAAAATGATGAAATTAAT 4500
Db 117059 TCTACTGTCACTTTATGAGAAATGTTGCAATTTGTTAAAAATGATGAAATTAAT 117000
Qy 4501 GTAAATTTAAAAAGACATATGATTTTGTGTAATTAAGTTGGCTGATCTAATAA 4560
Db 116999 GTAAATTTAAAAAGACATATGATTTTGTGTAATTAAGTTGGCTGATCTAATAA 116940
Qy 4561 GACATGAGAGAAATATCTTAAACAGAAAGTATGTTGTCTGTGGTACTAGATT 4620
Db 116939 GACATGAGAGAAATATCTTAAACAGAAAGTATGTTGTCTGTGGTACTAGATT 116880
Qy 4621 CTGAATCTCAGATTTCAACAACTACAGAGAGAACTTTCCAAATTAAGAGTGGG 4680
Db 116879 CTGAATCTCAGATTTCAACAACTACAGAGAGAACTTTCCAAATTAAGAGTGGG 116820
Qy 4681 GAGTGTGTATGACTGAACAGATCAAACTGTATTTCTTGTCTATTTCTGAAAAA 4740
Db 116819 GAGTGTGTATGACTGAACAGATCAAACTGTATTTCTTGTCTATTTCTGAAAAA 116760
Qy 4741 CTGCAATATPAACAGACTTATATAGCAATTTGATTTGTCACTTATTAATTAATT 4800
Db 116759 CTGCAATATPAACAGACTTATATAGCAATTTGATTTGTCACTTATTAATTAATT 116700
Qy 4801 AAATGATTAATGTATCTGGGAGAAAGTCAATAGATATACAAATACATATTAAGG 4860
Db 116699 AAATGATTAATGTATCTGGGAGAAAGTCAATAGATATACAAATACATATTAAGG 116640
Qy 4861 AAATGAGCATCTGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
Db 116639 AAATGAGCATCTGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 116580
Qy 4921 ACACAAAAATGACACTTGTGAGATTTGAATCTGAAGAGCTCAAGATCATATCAGAA 4980
Db 116579 ACACAAAAATGACACTTGTGAGATTTGAATCTGAAGAGCTCAAGATCATATCAGAA 116520
Qy 4981 TTCCAAAAATGT 5040
Db 116519 TTCCAAAAATGT 116460
Qy 5041 GGTTCACAGAGCATTAAGATTCAGTCTTATGAAAAAGAGAGAGAGAGAGAGAGCT 5100
Db 116459 GGTTCACAGAGCATTAAGATTCAGTCTTATGAAAAAGAGAGAGAGAGAGAGAGCT 116400
Qy 5101 TTGCTCTTCTATTAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160
Db 116399 TTGCTCTTCTATTAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 116340

Qy	5161	ACTTAAAGCACTGCTCATGGGATGTCATCCAGATTCAGAAAACCTGGATGTGGTC	5220
Db	116339	ACTTAAAGCACTGCTCATGGGATGTCATCCAGATTCAGAAAACCTGGATGTGGTC	116280
Qy	5221	CTTGTGCTGCTTACATTCCTAGAAAGTATGTCACAAAAGAAATCTTGGGGCA	5280
Db	116279	CTTGTGCTGCTTACATTCCTAGAAAGTATGTCACAAAAGAAATCTTGGGGCA	116220
Qy	5281	ACGAGAGTCTCTTCAAGCCCTGACCTGTCCTGATTCCTGCTCAATCAAGATTTTCA	5340
Db	116219	ACGAGAGTCTCTTCAAGCCCTGACCTGTCCTGATTCCTGCTCAATCAAGATTTTCA	116160
Qy	5341	GGAACTCTCAGAAATATAATGTCGGGAGAGAAACAGAACTGGAGTCTCGTCAGGA	5400
Db	116159	GGAACTCTCAGAAATATAATGTCGGGAGAGAAACAGAACTGGAGTCTCGTCAGGA	116100
Qy	5401	CTCAGGGACCAAGGGGCTGATTCAGACCTGCTCTTCACTGTTGTAACAGAAAACCT	5460
Db	116099	CTCAGGGACCAAGGGGCTGATTCAGACCTGCTCTTCACTGTTGTAACAGAAAACCT	116040
Qy	5461	TTAAATCTTAGGCTTAGCTTCACTTATGATTAATGAGATAATACATAGACAGTCT	5520
Db	116039	TTAAATCTTAGGCTTAGCTTCACTTATGATTAATGAGATAATACATAGACAGTCT	115980
Qy	5521	TTAAAGAACTCATAGCATGTTTAAACAAATGCTTAAATGTTGGTATCCACAGTGA	5580
Db	115979	TTAAAGAACTCATAGCATGTTTAAACAAATGCTTAAATGTTGGTATCCACAGTGA	115920
Qy	5581	AGACAGCATATCTTACTCTCTTACGAGATCTTCGGGTTTATGAGAGAAACATATCA	5640
Db	115919	AGACAGCATATCTTACTCTCTTACGAGATCTTCGGGTTTATGAGAGAAACATATCA	115860
Qy	5641	TACCAATCTATAGATGACAAACAGTTAGTGTCTGATGTGGATTAACAGAGTTCTC	5700
Db	115859	TACCAATCTATAGATGACAAACAGTTAGTGTGTGATGTGGATTAACAGAGTTCTC	115800
Qy	5701	CTTTTCTCCCATTTCTTTTGGGCAATCAGAGCTGTGGCAGTTGTCTTCTTAAGAG	5760
Db	115799	CTTTTCTCCCATTTCTTTTGGGCAATCAGAGCTGTGGCAGTTGTCTTCTTAAGAG	115740
Qy	5761	AGCTCATGATGATGATGATCTCTGATGCTCTCTATATCTCCAGAGAGATGATC	5820
Db	115739	AGCTCATGATGATGATGATCTCTGATGCTCTCTATATCTCCAGAGAGATGATC	115680
Qy	5821	TTCTTTCACCTGAGAGCTCTGCCATGTCATCTTGGGATTCAGACAAACGTGG	5880
Db	115679	TTCTTTCACCTGAGAGCTCTGCCATGTCATCTTGGGATTCAGACAAACGTGG	115620
Qy	5881	CCCTGTATAGGCAAAAAGAACTCTGAAATTTGTTCTTAAATGGCAGCATCTCTTA	5940
Db	115619	CCCTGTATAGGCAAAAAGAACTCTGAAATTTGTTCTTAAATGGCAGCATCTCTTA	115560
Qy	5941	TTTTTCCCTTATTTTCAATTCCTTCTCATCTGAGATTTGTTAGGTTAATTTT	6000
Db	115559	TTTTTCCCTTATTTTCAATTCCTTCTCATCTGAGATTTGTTAGGTTAATTTT	115500
Qy	6001	TTTTTTCAGCCCAATTTTGAATGCTGCAATCTTGAATTTAATCTTGAATCACTCTCTA	6060
Db	115499	TTTTTTCAGCCCAATTTTGAATGCTGCAATCTTGAATTTAATCTTGAATCACTCTCTA	115440
Qy	6061	TTTACCCCTCTTAACTGATTAATTCAGACATATGTCGTGGGTCCAAAGGGCTGTGA	6120
Db	115439	TTTACCCCTCTTAACTGATTAATTCAGACATATGTCGTGGGTCCAAAGGGCTGTGA	115380
Qy	6121	AAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	6180
Db	115379	AAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	115320
Qy	6181	CCGAGCAAAATATATCTCTCTTATATCTTAAATTTTCAATATGCAATAACAGAAATA	6240
Db	115319	CCGAGCAAAATATATCTCTCTTATATCTTAAATTTTCAATATGCAATAACAGAAATA	115260
Qy	6241	TACTAATTTTAACTCTAGGCTCTGATGATTTAAAGGCAATACCTAATTAACGTG	6300
Db	115359	TACTAATTTTAACTCTAGGCTCTGATGATTTAAAGGCAATACCTAATTAACGTG	115200
Qy	6301	TCAATCAGCTGCTGTATATCTCCCAATTAGACCTAATCTCATTTCTCAGTTGAAATTT	6360
Db	115199	TCAATCAGCTGCTGTATATCTCCCAATTAGACCTAATCTCATTTCTCAGTTGAAATTT	115140
Qy	6361	GCATGAATATCTCTTTTAAACCCAAAGCCCTTACATTTCTCTAATTTTCAATGAGCT	6420
Db	115139	GCATGAATATCTCTTTTAAACCCAAAGCCCTTACATTTCTCTAATTTTCAATGAGCT	115080
Qy	6421	CTCTCATATCAAAAGTTTGCATCAACAAAGAAAGCTAACAAATCTCCGAAAGAG	6480
Db	115079	CTCTCATATCAAAAGTTTGCATCAACAAAGAAAGCTAACAAATCTCCGAAAGAG	115020
Qy	6481	AATGAATAGGTTTACATTTGATGATCTAGCAGACAACTAGTATCCCATACATAT	6540
Db	115019	AATGAATAGGTTTACATTTGATGATCTAGCAGACAACTAGTATCCCATACATAT	114960
Qy	6541	TCCCACTTCAATTAATCTGCTGAGTGGCACTAGGCTCACCTCACTTACTTTCTCT	6600
Db	114959	TCCCACTTCAATTAATCTGCTGAGTGGCACTAGGCTCACCTCACTTACTTTCTCT	114900
Qy	6601	CTGTCTAATGCTAGGAAATTAAGCTGACAGCCACACCTTACCGAAACATGTTGACAA	6660
Db	114899	CTGTCTAATGCTAGGAAATTAAGCTGACAGCCACACCTTACCGAAACATGTTGACAA	114840
Qy	6661	ATGCTTCAAGGGGTTTCGCAAAAGCCACATGATGCTTATTTTAAAGGCATAGTGT	6720
Db	114839	ATGCTTCAAGGGGTTTCGCAAAAGCCACATGATGCTTATTTTAAAGGCATAGTGT	114780
Qy	6721	CAATTTCAAGTTTGGGCACTCATATCAATCTTCTCAACACAGATAGAGCTGTCA	6780
Db	114779	CAATTTCAAGTTTGGGCACTCATATCAATCTTCTCAACACAGATAGAGCTGTCA	114720
Qy	6781	AATGAATTCGATGAATGAATTTTCTCATCTAATTAATGATGTTCTTAATGCT	6840
Db	114719	AATGAATTCGATGAATGAATTTTCTCATCTAATTAATGATGTTCTTAATGCT	114660
Qy	6841	TACATGCTGCTTCAATTTTATTTTCAATTTCAATCAATCTAATGAGCTT	6900
Db	114659	TACATGCTGCTTCAATTTTATTTTCAATTTCAATCAATCTAATGAGCTT	114600
Qy	6901	CTCAGCATGATCTCTTCAATGATGAAGCTTATGAAGACATTTGCTCTTATG	6960
Db	114599	CTCAGCATGATCTCTTCAATGATGAAGCTTATGAAGACATTTGCTCTTATG	114540
Qy	6961	GAATAGGCACTAGAGATTAATAATGTAATGTCCTGCTGCAATGATGACACACT	7020
Db	114539	GAATAGGCACTAGAGATTAATAATGTAATGTCCTGCTGCAATGATGACACACT	114480
Qy	7021	GAGTTATTTCTCACCCACAGGTCCTGCAATTTTCAACATCTTACGGAAGATCCCAT	7080
Db	114479	GAGTTATTTCTCACCCACAGGTCCTGCAATTTTCAACATCTTACGGAAGATCCCAT	114420
Qy	7081	TCTCTGCTTCAATATGATGATCTTTTCTGTCAGAGATGACAGTCTGCTCATG	7140
Db	114419	TCTCTGCTTCAATATGATGATCTTTTCTGTCAGAGATGACAGTCTGCTCATG	114360
Qy	7141	AGGTTTCACAAACCACTCTTGTGATGTAATCTTCCACCTGAGAGAAATTTTCA	7200
Db	114359	AGGTTTCACAAACCACTCTTGTGATGTAATCTTCCACCTGAGAGAAATTTTCA	114300
Qy	7201	CCGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	7260
Db	114299	CCGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	114240
Qy	7261	CTTATGTTATCCCATGTCAGACAGAACTTGTGTGAGAGAGAGGTCAGGCTTC	7320
Db	114239	CTTATGTTATCCCATGTCAGACAGAACTTGTGTGAGAGAGAGGTCAGGCTTC	114180
Qy	7321	AGAGTCAACAAAGCTGATTTTCAACTGATTTGAGAACCCCTTTTGAATGAGTGA	7380

QY	8461	ACCTTGAGCTGTCGGGAGCAGATGTGAGCAGTGAAGAAGAAACCTTGCCTGTGAC	8520
DB	113039	ACCTGAGCTGTGCGGGAAGACGATTGAGCAGTGAAGAAACCTTGCCTGTGAC	1129800
QY	8521	AGGACTTGTGAGACATGTGCTGCCTGCACACCTTACATATATGTGATTTTCTTAGCC	8580
DB	112979	AGGACTTGTGAGACATGTGCTGCCTGCACACCTTACATATATGTGATTTTCTTAGCC	1129200
QY	8581	TTCTGCTCAGAAATGTCTCAGAGTCCCC	8609
DB	112919	TTCTGCTCAGAAATGTCTCAGAGTCCCC	112891
RESULT 4			
LOCUS	AC023078	163718 bp	DNA linear PRI 23-APR-2002
DEFINITION	Homo sapiens chromosome 11, clone RP11-583P24, complete sequence.		
ACCESSION	AC023078		
VERSION	AC023078.9	GI:20198710	
KEYWORDS	HTG.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 163718)		
JOURNAL	Barren, B., Linton, L., Nusbaum, C. and Lander, E.		
REFERENCE	Homo sapiens chromosome 11, clone RP11-583P24		
AUTHORS	Unpublished		
	2 (bases 1 to 163718)		
	Barren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burtett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Darelle, K., Dewar, K., Domino, N., Doyle, M., Fennestor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gargyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Landers, J., Lehocsky, J., Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Margus, N., McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Morrow, J., Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Sudriamanian, A., Talamas, J., Teefaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.		
	Direct Submission		
	Submitted (08-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
	3 (bases 1 to 163718)		
	Barren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarato, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Darelle, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargyna, S., Gande, S., Gerd, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, J., Lenocsky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margus, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mleaga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, W., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strausen, N., Sudriamanian, A., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.		

TITLE	Direct Submission
JOURNAL	Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (base 1 to 163718)
AUTHORS	Birren, B., Linton, L., Nubbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, R., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hango, B., Horton, L., Hulme, W., Illise, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Laskocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Michova, T., Mleणा, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhahang, P., Pierre, N., Polara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talmage, J., Tesfaye, S., Theodore, J., Tophan, K., Travers, M., Travis, N., Trifillio, J., Vassiliev, H., Vei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (23-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Apr 19, 2002 this sequence version replaced gi:15706197. All repeats were identified using RepeatMasker: Smit, A. F. A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
FEATURES	----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L6254 Center clone name: 583_P_24 ----- Location/Qualifiers 1. 163718 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="11" /map="11" /clone="RP11-583F24" /clone_1lb="RP11-11 Human Male BAC" complement(1359..1483) repeat_region /rpt_family="WIR" 1789..1864 /rpt_family="L2" complement(2486..2593) repeat_region /rpt_family="WIR" 3132..3312 /rpt_family="L2" complement(3321..3328) repeat_region /rpt_family="WIR3" 3523..3709 /rpt_family="WLR1J" 3815..4164 /rpt_family="MER2" 4460..4922 /rpt_family="LTR33" 5867..7374 /rpt_family="L1PA4" complement(8866..8938) repeat_region /rpt_family="L1PB1" complement(9021..9603) repeat_region /rpt_family="L1PB1"

```

repeat_region 10764. .10796 /rpt_family="AT_rich"  

repeat_region complement(11072. .11270) /rpt_family="L2"  

repeat_region 11819. .12173 /rpt_family="MLT1B"  

repeat_region complement(12553. .12696) /rpt_family="L2"  

repeat_region 13423. .13787 /rpt_family="L2"  

repeat_region 13788. .13816 /rpt_family="LMC4a"  

repeat_region 13817. .13965 /rpt_family=" (TA) n"  

repeat_region 13966. .14276 /rpt_family="LMC4a"  

repeat_region 14277. .14341 /rpt_family="AluJo"  

repeat_region 14358. .14420 /rpt_family="L1MC4a"  

repeat_region 14586. .14659 /rpt_family=" (TA) n"  

repeat_region complement(14765. .15433) /rpt_family="AT_rich"  

repeat_region 15192. .15199 /rpt_family="L2"  

repeat_region 15437. .16450 /note="<30 qual SNGL region"  

repeat_region 15437. .16450 /rpt_family="L1MD2"  

repeat_region complement(16453. .16536) /rpt_family="L2"  

repeat_region 18608. .18633 /rpt_family="MIR"  

repeat_region complement(18931. .19622) /rpt_family=" (TGAG)n"  

repeat_region 19616. .20496 /rpt_family="L1PAB"  

repeat_region 20970. .21019 /rpt_family="L1P"  

repeat_region complement(22343. .22686) /rpt_family="L2"  

repeat_region complement(24730. .24895) /rpt_family="L1MB4"  

repeat_region 25216. .25285 /rpt_family="MIR"  

repeat_region 26057. .26142 /rpt_family="L2"  

repeat_region complement(26150. .26307) /rpt_family="L2"  

repeat_region 26385. .26640 /rpt_family="MIR3"  

repeat_region 26755. .27105 /rpt_family="MLT1J"  

repeat_region 27397. .27858 /rpt_family="MER2"  

repeat_region complement(28675. .34751) /rpt_family="L2"  

repeat_region 34756. .36869 /rpt_family="L1PAB"  

repeat_region 36870. .37829 /rpt_family="L2RS"  

repeat_region 37830. .39164 /rpt_family="L1PAB"  

repeat_region complement(39165. .39467) /rpt_family="L1PAB"  

repeat_region 39468. .40356 /rpt_family="AluSx"  

repeat_region 40350. .40798 /rpt_family="L1PAB"  

repeat_region 40872. .41052 /rpt_family="L1MD2"  

repeat_region 41052. .41052 /rpt_family="L1MD3"  


```

Matches 7537; Conservative 0; Mismatches 968; Indels 420; Gaps 57;
QY 1 TGTATGAAGCCATGTCCTTTATATCCAAAACGAGAAAGATATATCAAA--AAAGAA 58
Db 117203 TCTATGAAGCCAGTGTCAACCTTAATCAAAAACGAGAAAGATATATCAAA--AAAGAA 117144
QY 59 ACTATGAGCCATGACATGATGAATATATCATGACAGAAATCCCAACAAAATCTAGCTA 118
Db 117143 ACTATGAGCCATGACATGATGAATATATCATGACAGAAATCCCAACAAAATCTAGCTA 117084
QY 119 ACCCATCCACAGCATATCAAGAGATATCCACCATTTGTCAAGTGGTTTCATACCA 178
Db 117083 ACCCATCCACAGCATATCAAGAGATATCCACCATTTGTCAAGTGGTTTCATACCA 117024
QY 179 GGGTGC-GGATAGGTTAATATACAGATCAATTAATGTGATACATCAATTAACGAA 237
Db 117023 GGGTGCAGGATGGGTTAATATACAGATCAATTAATGTGATACATCAATTAACGAA 116964
QY 238 TTAATAAACAATAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAAATCT 297
Db 116963 TTAATAAACAATAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAAATCT 116904
QY 398 AACATTTCTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGACATACCTTAAT 357
Db 116903 AACATCCGATGATTAATAACCTTCAGCAAAATGCGCATAGAAAGACATATCTTAACA 116844
QY 358 TATATAAAGCCATATATGACGACCCACAGCAAACTATATCTGATGGGAAAAATGGA 417
Db 116843 TAGTAAAGCCACCGATGACAACTATATGCTTAATATGAGGAAAAATGGA 116784
QY 418 AAACATGTCCCTGAG-----AACTGGAACAAGACAAGATG-CTACTTTCACCACTT 469
Db 116783 AAACATCTCCAGAGTTGAAAAATCTTCCCAAGCAAAAGATGCCACTTTCACCACTT 116724
QY 470 CTATTCACATAGTATGGAAGTTTATGCGCAGCATCGACAGAAAGAAATCAAG 529
Db 116723 CTGTTCAACATAGTATGGAAGTTTATGCGCAGCATCGACAGAAAGAAATCAAG 116664
QY 530 GGCACCCCAATCAATTAAGAGAACTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
Db 116663 GGCATCCCAATCAATTAAGAGAACTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 116604
QY 590 TACCTAGAAAACCCCTAAGACATCATCGAAGAACTCTAGAACTGATACATTAATTCAGT 649
Db 116603 TACCTAGAAAACCCCTAAGACATCATCGAAGAACTCTAGAACTGATACATTAATTCAGT 116544
QY 650 AAAGTTTCAGATACAACTAAATGTACCAAACTAGTACCTGCTATACCAACAGT 709
Db 116543 AAAGTTTCAGGATACAACTAAATGTACCAAACTAGTACCTGCTATACCAACAGT 116486
QY 710 GACCAAGCTGAGATTCATTAAGAACTCAAACTTTTCAATAGCTGTAAAAA--- 766
Db 116485 GATCAAGCTGAGATTCATTAAGAACTCAAACTTTTCAATAGTGTAAAAA--- 116426
QY 767 -----TACTTAAGATATTTCTTAACCAAGAGGTGAAGACTCTTACAGAA 813
Db 116425 AAAAAAACAACCTTAAGATATTAACCTTAAGAGGTGAAGACTCTTACAGAA 116366
QY 814 AAACCTACAAACACAGCTGACATCATAGATGACAAACAAAGTGAACACATCCATGC 873
Db 116365 AA-----CTATGAAACCTGTAATCATATAGATGACAAACAAAGTGAACACATCCATGC 116309
QY 874 TCAATGATGGTGAATTCATATTTGTGAATAGCAACCTATTTGCCAAAAGAACTTCAAG 933
Db 116308 TCAATGATGGTGAATTCATATTTGTGAATAGCAACCTATTTGCCAAAAGAACTTCAAA 116249
QY 934 TCAATGATGGTGAATTCATATTTGTGAATAGCAACCTATTTGCCAAAAGAACTTCAAG 993
Db 116248 TCAATGATGGTGAATTCATATTTGTGAATAGCAACCTATTTGCCAAAAGAACTTCAAG 116190
QY 994 TCTAAATTCATATGGAACAACAACCAAAAAAATTTTCAAGAACTTCAAGAACTTCAAG 1053
Db 116189 TCTAAATTCATATGGAACAACAACCAAAAAAATTTTCAAGAACTTCAAGAACTTCAAG 116145

QY 1054 AGACTTAGCAAAAAGAAACAATCTGAGGCAATCACTTACCATCTTCAAACTATCTAG 1113
Db 116144 AGCTTAGCAAAAAGAAACAATCTGAGGCAATCACTTACCATCTTCAAACTATCTAG 116085
QY 1114 AAGCTATATATCAACCAAAACATCATGACCTGACATTAACCTAGGACATAGCAATGG 1173
Db 116084 AAGCTATATATCAACCAAAACATCATGACCTGACATTAACCTAGGACATAGCAATGG 116025
QY 1174 AAAGAAAGAAATTCAGAAATTAAGCAATTAATATATAGCAACCTGATTTTGACAA 1233
Db 116024 AAAGAAATGAAATTCAGAAATTAAGCAATTAATATATAGCAACCTGATTTTGACAA 115965
QY 1234 GCAACCAAAAACATTAAGTGGGAAAGACATTTAGTTAAACAATTTGTGATGATAT 1293
Db 115964 GCAACCAAAAACATTAAGTGGGAAAGACATTTAGTTAAACAATTTGTGATGATAT 115905
QY 1294 TGGCAAGCCACATGAGAAATGAAATGAACTGATCCCTGTCTCTGACTTAATACAAAAT 1353
Db 115904 TGGCAAGCCACATGAGAAATGAAATGAACTGATCCCTGTCTCTGACTTAATACAAAAT 115846
QY 1354 TGTATCAAGATGATCAAAAGACTTAATTTGTGACCTTAACCAATTAATTTTGAA 1413
Db 115845 CAACCTCAAGATGATCAAAAGACTTAATTTAGCACTTAAGCAATTAATTTTGAA 115786
QY 1414 TAAATCAGAAAAATGCTTCTAGACATTTCACTTGAAGCAAGCTTCACTGCAAGAACCC 1473
Db 115785 TAAATCAGAAAAATGCTTCTAGACATTTCACTTGAAGCAAGCTTCACTGCAAGAACCC 115726
QY 1474 AAAGTAAATGCAACCAAAAACCAAAATTAATATAGATAGACCTTAATTAACCTAAAGCTT 1533
Db 115725 AAAGTAAATGCAACCAAAAACCAAAATTAATATAGATAGACCTTAATTAACCTAAAGCTT 115666
QY 1534 TTGCGAGCAAAAACATATATATGACAGCAACCAAGCAACCAAGCTGAGAGAAAT 1593
Db 115665 TTGCGAGCAAAAACATATATATGACAGCAACCAAGCAACCAAGCTGAGAGAAAT 115606
QY 1594 CTTCAAACTAAGCACTGACCTAAGAACTAATATCCGAAATCCCAAGAACTCAACA 1653
Db 115605 CTTCAAACTAAGCACTGACCTAAGAACTAATATCCGAAATCCCAAGAACTCAACA 115546
QY 1654 AATCAGCAAGAAAGAAACAACATCCATGAAAGTGGGCTTAAGACATGAATAGACA 1713
Db 115545 AATCAGCAAGAAAGAAACAACATCCATGAAAGTGGGCTTAAGACATGAATAGACA 115490
QY 1714 ATTCTCAAAAGAAATATCAAAATGCGCAACAAAC--AGAAAAATGTTTAACCTACT 1771
Db 115489 ATTCTCAAAAGAAATATCAAAATGCGCAACAAACATGAAGAAATGTTTAACCTACT 115430
QY 1772 AATGATTAGGAAATGTAATCAACCTGAATGCAATCACTTAACCTGCAAGAAAT 1831
Db 115429 AATGATTAGGAAATGTAATCAACCTGAATGCAATCACTTAACCTGCAAGAAAT 115370
QY 1832 GGTATATTTTAAATCTTAAATTAATATGATGTTGGTGTCTGTGTGATTAAGAAC 1891
Db 115369 GGTATATTTTAAATCTTAAATTAATATGATGTTGGTGTCTGTGTGATTAAGAAC 115310
QY 1892 ACTTTTACCTGCTGTGGGAAATGTAATCTTCCGCAACCACTATGAAACAGTGTGAA 1951
Db 115309 ACTTTTACCTGCTGTGGGAAATGTAATCTTCCGCAACCACTATGAAACAGTGTGAA 115250
QY 1952 ATTCTTAAAGAACTAAAGTATGACATTTTGTGATCCAGAACTCCCA----- 2000
Db 115249 ATTCTTAAAGAACTAAAGTATGATCAATTTTGTGATCCAGAACTCCCA----- 115190
QY 2001 ----- 2000
Db 115189 CTACCAAGAGAAAGAAAGTCAATTTTACGAAAAAGCACTTGCATGTGATTTATAGC 115130
QY 2001 ----- 2000
Db 115129 AGCACAATTCATTAACAAAAATATAGAAACAGCTCAAAATGCCATCAATCAACAAAGTGG 115070

OY	2001	-----		2000
Dd	115069	ATAAAGAAAATTATATATATATAGTGTGTAATGTAATATGTATATATATATATATATA	115010	
OY	2001	-----TTAATATATATATATATATATATATTA-----TATACATGG	2037	
Dd	115009	TGTGCTGTATATATATTTATATATATATATATATATATACACACACACACATATATCATGG	114950	
OY	2038	AATACAACACTCAGCCATTAATAAAAAAATAAAAATGAATGCATTCACAGCAATCTTAGATGGAAT	2097	
Dd	114949	AATACTACTCAGTAATATAAAAAGAAATGAATATATATAGCATTTCTCAGCAACCAGATGGAAT	114890	
OY	2058	TGGAGAACCCTTATTTCTAAATGGGGTATCTCAGGAATGGAAAAACCAACATCATATGTTCT	2157	
Dd	114889	TGGAGAACCAATTTATTTCTAAATGGAAGTAACACAGAAATGGAAMAAACCAACATCATATGTTCT	114830	
OY	2158	CACCTTACAAGTGSGGGGCTAAGCGTGAAGACACGAAGGAGATGAATATATATGAATCTC	2217	
Dd	114829	CACCTTATATATGGAGCTAATGCTGTGAAGACACAAAGGCCATGAAATATATATCATTAACCTT	114770	
OY	2218	TGGGACCTTGAAGGGAGAGATGGAAGAGAGCGAGGGATTAAGAATACACATATGGGTAC	2277	
Dd	114769	TGAGGACCTTGGGGGGATGGAATGAAGAGAGATGAAGATTAAGAATACACATATGGGTAC	114710	
OY	2278	AGTGTATCACTGCTCAGGTATATGGGTATCACCAAATCTCGAATATTCACATAAGAACTT	2337	
Dd	114709	AGTATATACACTGCTTGTGGGTATGGGTATGCACAAATCTCGAATATTCACATAAGAACTT	114650	
OY	2338	ATCCATGGAAGCAACACCACTGTCCTCCCAAAATCCCAGATGAATTAATTAATTA	2397	
Dd	114649	ATACATGTATACAAACACCACTGTCCTCCCAAAATCCCAGATGAATTAATTAATTA	114593	
OY	2398	TAAATGATTTATTTTACAGAA--TTTTAAAAAGTTCACTGTTCAGAGTTTATATATATGA	2456	
Dd	114592	TAAATGATTTGATTTTACAGAACTTAAAAAAAGTTAGCTGTTTCAAGTTGATTA--AAAG	114535	
OY	2457	AGTAAGAATGAAGAAGTATGCAAGTGTATGCTCTGGAACAATGGACTCTAGATTTTCAC	2516	
Dd	114534	AGTAAGAATGAACAGTGTATGCAAGTGTATGCTCTGGAACAATGGAGTTCTAGACTTTTCAC	114757	
OY	2517	CTTGATATCACTTCTCTGSCATTTTGGAAAAAGTATACATNGAATATATATATCAACACTAT	2576	
Dd	114474	TTTGATATCACTTCTCTGSCATTTTGGAAAAAGTATATACATNGAATATATATATCAACTAT	114415	
OY	2577	GATTAAGAAAAATC--AAAAAATGTGTACAGGCATTTGTACAGCCTTGAATGTCCCAT	2634	
Dd	114414	AATTAAGAAAAACCAAAAAAATGTGTGTACAGCATTTGTACAGCCTTGAATGTCTCATG	114355	
OY	2635	ATCTAATTTTTCAATTTGGATTAAGCTCATATATATATGTTCAATGCTTAATGTAT	2694	
Dd	114354	ATCCACTTTTTTCAATTAATTTATTAAGCTCATATATATATGTTCAATTAATGATGTAT	114295	
OY	2695	GCTTATGCCCCATTAATTTGATTTTTTAAATATCAGACACAGCAATTTACTTACAGAAATTTGAACA	2754	
Dd	114294	ACCTTATGCCCCATTAATTTGATTTTTTAAATATCAGACACAGCAATTTACTTACAGAAAGTTGAACA	114235	
OY	2755	AGATGGAGCTGATATAGAGAGGCTTAAATGTATCTGATATATGGAACAAGGCCCAAGATCA	2814	
Dd	114234	ACATGGAGCTGATATAGAGAGGCTTATATGTACTGATATATGGAACAAGGCCCAAGATCA	114175	
OY	2815	TCTCAGTTAGATTGTGTCTCAAAATACCTTGAGCCTCTGATTTTGCCCATATGTTCTCAT	2874	
Dd	114174	TCTCAGTTAGATTGTGTCTGAACTTCTGAGCCTCTGATTTTGTCCACAGTCTCATG	114115	
OY	2875	CAGGAATATPACAAGACTGTTCACAGCACTTCTGTATGAGCTGATTTGTCTCACAGCTTTCATT	2934	
Dd	114114	CAGGAATATPACAAGCTGTTCACAGCACTTCTATATAGCTGATTTGTCTCACAGCTTTCATT	114055	
OY	2935	TCAGCTCTGTATGAGCATCTCTGAAATTAAGCAACAGAAAAAGTCTCTGAAATGTCAGTA	2994	
Dd	114054	CTAGCTCTCCGTATGAGCATCCCTCTGAATTAAGAAACAAAAAGGTCCTCTGAATGTACTGA	113995	
OY	2995	ATCCAGAAAGGCTCTCTAATCTTTAGACACAGGAGGTCTTCAACAATGACA---AAG	3050	

Db	113994	ATCCGAGAAAGGCCCTCCACCTTTACCAACAGGAGATCTTCACTCTGACAAAAGAA	113935
Qy	3051	AAGAAACGATTAAGGGTAAGTACCAAGAACTCTTTCTTCCACAGTCAGTTAATTTTG	3110
Db	113994	AAGAAACGATTAAGGGTAAGTACCAAGAACTCTTTCTTCCACAGTCAGTTAATTTTG	113875
Qy	3111	CTGTAAATCATATGTCCTTAATGCTTCCACTTGGTGTACATGAGGGGTACAGACTTG	3170
Db	113874	CTACAAAGATCACTTCTCGTGTTCACCTCCGTCTACATGAGGGGATCATGATCTTG	113815
Qy	3171	TTTTCAGAAAAGACAGAGACATGAGCTTCCTTACAGAACTGAGTGTCAACCCAA	3230
Db	113814	TTTCATGAAAAGACAGAGACATGAGCTTCCTTCTGAAACCGATGTGCTACAGCCAG	113755
Qy	3231	ACTGTGAGCTCTTAATGAGTGTCCGCCCTTCAATTTATCTCCCATATCACTCTTC	3290
Db	113754	TCTGTGTAGTTCCAAATAGAGTCCCGCTCTGTATTTATCTCCCATATCACTCTTT	113695
Qy	3291	ATTCGAATCATTAATCTGCTCCCTCAGAGAGACTGCTGCTTTAATTCAATTTAACA	3350
Db	113694	ATTCGAATCATTAATCTGCTCCCTCAGAGAGAAATGCTAATCTTAATTTAATGA	113635
Qy	3351	GCAAGGGACATGACGCAATTTCTTCCAGAGTTAACTGCTATAG - ACCAGTTCTTT	3409
Db	113634	GCAAGGGACATGACGCGCTTGCTTCTAGACTGAACGTGCTGAAACCCAGTTTGCT	113575
Qy	3410	GTTTCACTTAATTTCAATTTAATTTCTTTGCTATCTGAGAAAGGTAAAGAAAGATA	3469
Db	113574	GTTTACTTACTTTCAAAATTTTCTTCTTCTGCTATCTGAAAGTTTCAAGAAAGACA	113515
Qy	3470	TAGATGCGCC-----AATAATTAAGAGTGTTCATGAGGAAAGTATTTACAAAGATG	3522
Db	113514	TGGAATGTCCAAACTGAATATGATGAGAGTGTTCATGAGGAAAGTATTTACAAAGACA	113455
Qy	3523	CACAGAGTTAA-----GGGTACAGATCTTAAGCAAGCAATCATAGGGG	3565
Db	113454	TGACAGGGTTAAAGGAAAGAAATGATCATGGGTACACACCTTAAGCAATCATAGGGG	113395
Qy	3566	AGCACTACTTCCCTCCCTTAGGCTGAAGCGAAGGGAAGGAGACAGTTACATGTGCT----	3621
Db	113394	AGCACTTAAGTCTCTCCCTTAGACTGAAGGAGACAGGAAAGGACAGTTATTAATGTGCACA	113335
Qy	3622	--GCCATAGCCATAGCTGTAGCCATTAAGGGGTGGAGAGCATGAGACGCAAGTGAAGAAG	3679
Db	113334	TAGCATATAGCCATATCTGTATTCATATAGGGGTGGAGAACAAAGCGGGCAGGTGGAGAAAG	113275
Qy	3680	CCCTGCGTGGCCACGACACAGCCACACAGGCTGATATAGTTGGATCTGTGTTCCACCA	3739
Db	113274	CCCTGTATAGCCACACACAGCCACACAGGCTGATATAGTTGGATCTGTGTTCCACCC	113215
Qy	3740	AAATCTCATGTGATTTGAATTTCCAAATGTGGAGAAAGGGCTGTGGGAATGATTAAT	3799
Db	113214	AAATCTCATGTGATTTGAATTTCCGAATGTGGAGAAAGGGCTGTGGGAATGATTAAT	113155
Qy	3800	TAGATCAGCGGGATGTTTGCATGAATGTTTAAACACATCCCTTGGTATGTTGT	3859
Db	113154	TAGATCAGCGGGATGTTTTCATATGAATGTTTAAACACATCCCTTGGTATGTTGT	113095
Qy	3860	TGTGATCTGACGAGTTCTCATGAATCTAAGTTTAAAGCGTGTAGCACTCTCTCC	3918
Db	113094	TGTGATCTGATGAATTTCTCATGATATCTAAGTTTAAAGCGTGTAGCACTCTCTCC	113035
Qy	3919	-CTGCTCTTACTCGTGCTTCACATGTGAGAGGCTCGCTCCCTTGGCTTTACAC	3977
Db	113034	TCTCTCTTGTCTGTGCTTTCACATGTGAGATGCTCTCTCTCCCTTGGCTTTACAC	112975
Qy	3978	AGATATGAAGCTTCTCTGAGGCTTCCACAGAGCAAGAACTGCTATGCTTTGTACAGT	4037
Db	112974	ATGATATGAAGCTTCTCTGAGGCTTCTTAAGAAAGCAAGAGCTGTATGCTTCCGTATAG	112915
Qy	4038	CTGTAGAGCTATTACCGCATTTAACCAATTTCTTCAATTAATTTCCCACTCAGGTATTT	4097

Db	112914	CTGCAGAGCCATGAAACCAATTAAACCTGTTTCTTATTAATTTCCCACTTCAGGATAT	112855
Oy	4098	TCTTTTATGCAATTTGAGATGAATCTAATATACAGACAGAGAGCCAGAGATGGAATCC	4157
Db	112854	TCTTTGTAGCAATTTCAAGATGAACTTAATATACCTGGCAGAGAGCCAGAGGTGGAATCC	112795
Oy	4158	CAAGGTGCTTCTCTGTCTTCCAGTCTTCCTGTGTGTCTTCCAGTGTCTCAATTTCCA	4217
Db	112794	CAAGGTGCTCTCTGTGTCTTCCAGTCTTCCTGTGTGTGTCTTCCAGTGTCTCAATTTCCA	112735
Oy	4218	CCAGAAACCAAGAAATTTAAAGAATCCCACTGATGTGTGTCATGTGAAGCCACTCTTGGG	4277
Db	112734	CCAGAAACCAAGAAATTTAAAGAATCCCACTGATGTGTGTCATGTGAAGCCACTCTTGGG	112675
Oy	4278	ATGTCAAAACAGATTAAGAAGAATGGAAGCAAAATCCTCAAGGTAAATGAGATACCT	4337
Db	112674	ATGTCAAAACAGATTAAGAAGAATGGAAGCAAAATCCTCAAGGTAAATGAGATACCT	112616
Oy	4338	CTCACCTTCTTGTATCTCTCTAATCTCTGGGGCTTCTCTATC-----TGATTAATCC	4391
Db	112615	CACACCTTCTTGTATCTCTCTAATCTCTGGGGCTTCTCTATC-----TGATTAATCC	112556
Oy	4392	TGTTCATTTGAGCTCTATCAGACCTAATTAAATGTTGGCTGTCTTCTCTAATCTCAG	4451
Db	112555	GCTTCATTTGAGCTCTATCAGACCTAATTAAATGTTGGCTGTCTTCTCTAATCTCAG	112501
Oy	4452	TTTATGCAAAATGTTGTCATTTGTTTAAATATGATAGAAATTAATATGTAATTTTAA	4511
Db	112500	TTCTATGCAACATGTTTGCGTTCAATCAAAATATGCGTAAGAAATTAATATGTAATTTTAA	112441
Oy	4512	AAGA-----ACATATGATTTTGTGTTGAATATTAAGTTTGCTATCTAATTAAGACATGA	4567
Db	112440	AAGAGAGGACATATTTCTTTTGTGTTGAATATTAAGTTTGCTATCTAATTAAGACATGA	112381
Oy	4568	AGAAGAAATATCTTAAACAAGAAAGATATGTTGTGCTGCGGTCACTAGGTTCTGAATC	4627
Db	112380	AGTAAATATATCTT--AACAAGAAAGATATGTTGTGCTGCGGTCACTAGGTTCTGAATC	112322
Oy	4628	TACAGATTCACAACTACAGAGAGAACTTTTCCAAAATTAAGGTGTGTGCGAGTTGT	4687
Db	112321	TGCAATTTCAACAAACACAGGTGGAACATTTTAAAAATTAAGATATGTGGCGAGTTGT	112262
Oy	4688	GTAATGACTGAACAGGTACAACTTGTATTTCTTTGTCATTAATTTCTGAAAACTTCAAT	4747
Db	112261	GTAATGACTGAGCAGGTATTAACCTTGTATTTCTTTGTCATTAATTTCTGAAATGATGAGT	112202
Oy	4748	-ATTAACAACAATTATATAGCAATTCATTTGTCTCAAGTATTTCTAAATACTT--TAATG	4805
Db	112201	AAAAAAAACAATTTACATAGCAATTAATTTGTCTCAAGTATTTCTAAATACTTAAATAAT	112142
Oy	4806	ATTTAATGATCTGGAGAAAGTCATAGATATATACAATATCCATAT-----ATTA	4858
Db	112141	ATTTAATGATCTAGAGAAAGTCATAGATATATACAATATCAATGTGTTTGTATA	112082
Oy	4859	GGAATTTGACATCTGCAATTTTGTCTGTGTGTGGGTTCTGAAAAGATCCCTGTGTA	4918
Db	112081	AAGAATTTGACATCTGCAATTTTGTGTGTGTGTGGGTTCTGAAAAGATCCCTGTGTA	112022
Oy	4919	ATATCAAAAAATGACACTCTTGCAGATCTGAATCTAAGAGTCCAAAGCATCATATCAG	4978
Db	112021	ATATCAAAAAATGACACTCTTGCAGATCTGAATCTAAGAGTCCAAAGCATCATATATCAG	111962
Oy	4979	AATTCAAAAAATGTGTGTCTCCAGATTCCTAGAGAGTGTGCTCATCTGTGTATCTTAC	5038
Db	111961	AATTC--AAAAATTTGAATCTCCAGAGTGTCTAGAGAGTGTGCTCATTTTGTGTATCTGTG	111903
Oy	5039	ATGTATTCAGAGCACAATTAGCATTTCAAGTCTTAAGAAAAAGACAGAGGGAAGAGAGG	5098
Db	111902	ATGATTTCCAGACCAATCGCATTTCCAGCATCATATG--AAAAAGATGAAGAAAGAGAGG	111844
Oy	5099	CTTTGTCTCTCTTATTAATCTCATAGAGCCAGAGACTGCTTGTCTCATTTTGTGATTTCT	5158
Db	111843	CTTTGTCTCTCTTATTAATCTCATAGAGCCAGG-----GATCATTTTGTGCACTTT	111793

QY	5159	CCACTTAACAGACCCCTGCTCATTGGGATGTCAATCCAGCATCAAGGAAAATCTGGGATGTGGG	5218
Db	111792	ACATTAAGAGCTCTGCTCATATATATAGCA--CAATCATCAAGGAACACTGGGATGTGGG	111734
QY	5319	TCCATTGTGCTGTGTACATCTCCAGAAAGTTATGTGACCMAAAAGAAATCTT-----	5274
QY	5275	-GGGGCAACGAGAGCTCTTCAAGCCCTGAGCTGTCTCTGATTTCTGTGGCTCACTAAGA	5333
Db	111673	GGGGCAACCTGAGAGCTCTTCAAGCCCGGCACTGTCTCTGATTTCTGTGGCTCACTAAGA	111614
QY	5334	TTTTTCAGGAATCTCTCAGAAATTAATTAATGATGGGGCAGAGAACAGAACTGAGTCTCG	5393
Db	111613	TTTTTCAGGAATCTCTCAGAAATTAATTAATGATGGGACAGAGAACAGATTCAGAAAGTCTCA	111554
QY	5394	TGACAGACTTCAGAGAACAGAGGGCTGTGATTTGGAACCTGTCTCTTCAATGTTGTGAAC--CAGG	5452
Db	111553	AACAGGGCTCCAGAGGACAGAGGACGTGATTTGGAACCTGTCTCTTCAATGATTTGATCTCAGG	111494
QY	5453	AAAAACCTTTAATTTCTAAGGCTTAGGCTCATCTTAATGATTAATGAGTAATATACATA	5512
Db	111493	AAAAACCTTTAATTTCTAATGAGGCTGAGCTTCACTTAATGATTAATGAGTAATATACATA	111434
QY	5513	GACAGCTTTTAAAGAACATCATAGCATGTTTAAACAACATGCTAAATGTTGTGATACAC	5572
Db	111433	GATGATCTTTAA-----GAAAGACATGCTAAATGTTGTGATGATCTAC	111392
QY	5573	AGTGAAGAAAAGACGAGCATGATCACTTACCTTTACCGATCTTCCGGGTTTATAGAGAAAGCA	5632
Db	111391	AGTGAAGAAAAGACG-----CTTAATGAGCTTCAAGGTTTCAATGAGAAAGTTAA	111345
QY	5633	ACATATCATATCCATATCCATTAATGATGAGACAACAAGTTTATGTCTCTGAGTGTGATTAACAG	5692
Db	111344	ATATATCATATCCATTAATTAATGATGAGACAAGTGTTTATGTCTCTGATGATGATTAACAG	111285
QY	5693	AGGTTCTCTTTTCTCCCATATTTCTTTTGGGCCAATCAGAGCTGTGGCAGCTTGTCTC	5752
Db	111284	ATGTTCTCTTTTCTCCCATATTTCTCTTTCCGGGCCAATCAGATTAATGAGCAGCTTGTCTT	111235
QY	5753	CCTAAGAGAGCTCATGATGATGATGACTCACTCTCATGCTCTCTATACTCCAGAGAG	5812
Db	111224	TCTACTAAGAGTACGATGATGATGATACATTAATTTCTTATGTCTCTGTACTCCACAGAG	111555
QY	5813	GATGATCTTCTTCCACCTGGAAGAGCTCTGCCCATGTGATTTCTTGGATTCAGAGC	5872
Db	111164	GATGAATCTTCTTCCACATGGAAGAGCTCTGCCATGTGATTTCTTGGGATTCAGAGC	111105
QY	5873	AAACGTGGCTCTGATAGGCAAAAAGAACTCTGTAATTTGTTCTTAAATGGACGCACT	5932
Db	111104	AAACGTGGCTCTGATAGGCAAAAAGAACTCTCTAATTTGTTCTTAAATGGATTTTCACT	111045
QY	5933	CACCTCAT--TTTTCCCTTATTTCAATTGCTTCTCATTTCTCATCTGAGAGTTGTTTAAAG	5991
Db	111044	CTCCTCATATTTGTTCCCTAATTTGATTTACTTCTCATTTCTCATCTGAGAGTTGTTTAAAG	110985
QY	5992	TTAATTTTTTTTTTCAAGCCACAATTTTGACGTGCAACTTGGATTTAATTGAGAAATCAC	6051
Db	110984	TTAATTTTTTTTTTCAAGCCACAATTTTGACGTGCAACTTGGATTTTCACTTGAAATCAC	110925
QY	6052	TTCCTTACTTTACCCCTCTTAACTGTATTAATGACAATAGTGGTCTGGGTCCAAAG	6111
Db	110924	TCTGTGTCTTTA--CTCCCTCTAATGATTAATTCACACAGAGTCTGTCTGTTCGAAAG	110866
QY	6112	GGCTGTGAAAAAATGATCATGAGTCAAGCCCTGTGGGCTCACATTTATATATATAT	6171
Db	110865	GGCTAGTCAAAAGATCGATCATGAGTCAAGCCCTGTGGGCTCAACTTATATATC-----T	110813
QY	6172	ATATATACCCCGGACAATATATATCTCTTTATCTATATTTCAATTTATCTGCATA	6231
Db	110812	ATATATATCTCCAGACAATATATATCTTTTGTGTCTC--CAGGCAATTAATGGCCATA	110755

QY 6232 CAGGAATATACTAATTTTAACTCTAGGCTCTTCAGATGATTTAAAGAGCAATACCT 6291
 DB 110753 TGGGAATAAATCACTATTCCTAGGCTCTTCAGATGATTTAAAGAGCAATACCT 110694
 QY 6292 AATAAATCTCAATCAGCTGCTGTTATTTCTCCAAATAGACCTAATCTCATTCCTCAG 6351
 DB 110693 AATPAATCTCAATCAGCTGCTGTTATTTCTCCAAATAGACCTAATCTCATTCCTCAG 110634
 QY 6352 TTGAATTTTGAGAAATATCTCTCTTTTCAACCCAGCCCTTACCTTCTCTATTTTCAC 6411
 DB 110633 CTGAATCTGCAATCTCTCTCTTATATCTCAAGCCCTATCTAGCCCTATTTCC-C 110575
 QY 6412 TCATGACTCCTCTCATACCAATGTTTGCATCAACCAAGAGCTAACCAAGATCTCCC 6471
 DB 110574 TCATGACTCCTCTCATACCAATGTTTGCATCAACCAAGAGCTAACCAAGATCTACT 110516
 QY 6472 GAAAGAGAGAAATGAGTTTACATTTGTATCTCAGAGAACTTAGTAGTCCCC 6531
 DB 110515 GAAAGAGAGAGAAATGAGTTTACATTTGTATCTCAGAGAACTTAGTAGTCCCC 110456
 QY 6532 CATACATATTTCCCACTTCAATTAACCTGCTGAGTGCATCAGGCTCACTCACTTA 6591
 DB 110455 CATGATATTTCCCACTTCAATTAACCTGCTGAGTGCATCAGGCTCACTCACTTA 110396
 QY 6592 CTCTTCTCTGTTCTATGCTGAGCAATTAAGCTCAGACCCCACTTACCCAACT 6651
 DB 110395 CTCTTCTCTGTTCTATGCTGAGCAATTAAGCTCAGACCCCACTTACCCAACT 110336
 QY 6652 GTGATC-AAAATGCTTAGGGGTTTGGGCAAGCCAGCTAGTCTTATTTTAAAGCA 6710
 DB 110335 GTGATC-AAAATGCTTAGGGGTTTGGGCAAGCCAGCTAGTCTTATTTTAAAGCA 110276
 QY 6711 CATCAGTGTCAATTTAGGTTTGGGCACTCATCAATCTTCTTCAACAGAGATGA 6770
 DB 110275 CATCAGTGTCAATTTAGGTTTGGGCACTCATCAATCTTCTTCAACAGAGATGA 110216
 QY 6771 GCTGTCCAAATGAAATCTGATGAATGAATTTTCTTCACTAATTAATGCTGTGT 6830
 DB 110215 GCTGTCCAAATGAAATCTGATGAATGAATTTTCTTCACTAATTAATGCTGTGT 110156
 QY 6831 TCTAATGCTTCAATGCTGCTTCAATTTTATTTTCCATTCATCAATCTTACCATTCG 6890
 DB 110155 TCTAATGCTTCAATGCTGCTTCAATTTTATTTTCCATTCATTCATCAATCTTACCATTCG 110096
 QY 6891 CATTAGGCTTCTCAGTGAATCTTCAATGAATGAATCTTATGAAGCACTTGT 6950
 DB 110095 CATTAGGCTTCTCAGTGAATCTTCAATGAATGAATCTTATGAAGCACTTGT 110036
 QY 6951 GCTGCTTATGG-ATAAGCACTAGAGATTAATTAATGTAATGCTGTCTGCAATG 7009
 DB 110035 GCTGCTTATGG-ATAAGCACTAGAGATTAATTAATGTAATGCTGTCTGCAATG 109976
 QY 7010 ACTGACACACTGAGTTATTTCTACCCAGCCAGGCTCCGCCATTTTCAAC-ATCCTAGC 7067
 DB 109975 ACTGACACACTGAGTTATTTCTACCCAGCCAGGCTCCGCCATTTTCAAC-ATCCTAGC 109916
 QY 7067 GAAGATCCATTTTCTCTGCTTCAATGATGATCTTTTCTGCTGCTGAGATGATGCC 7127
 DB 109915 AAAGGTCTCATTTTCTCTGCTTCAATGATGATGATCTTTTCTGCTGCTGAGATGATGCC 109856
 QY 7128 AGTCTGCTCATGAGGCTGTCACACCACTCTTTGTGTATCTGAATTTCTTCCACTGAG 7187
 DB 109855 AGTCTGCTCATGAGGCTGTCACACCACTCTTTGTGTATCTGAATTTCTTCCACTGAG 109796
 QY 7188 AGAATAATTTGAGGCGGAGATAGATATCATGGGCTTCCAGAGCTGGCTAGATGATG 7247
 DB 109795 AG-AAAATTTGAGGCGGAGATAGATATCATGGGCTTCCAGAGCTGGCTAGATG 109737
 QY 7248 GGGGCTGTTTGTATCTTATGTTATCCCATGTCAGACAGAACTTGTGTGCTAGATGA 7307
 DB 109736 GGGGCTGTTTGTATCTTATGTTATCCCATGTCAGATGACAGAACTTGTGTGCTAGATGA 109677
 QY 7308 GAGGTACGGCTTCAAGATCAACAGAACTGATTTCAAACTGATTTGAGGACCCCAACC 7367

DB 109676 GAGGTACGGCTTCAAGATCAAGAGAACTGATTTTCAAA-TGGAATTTGAGGACCCCAACC 109618
 QY 7368 TTTTGAATGAGTATTTATCTGTCGAGTCTCTG-ATCTCTCTCTTT-AAATGAGACA 7425
 DB 109617 TTTTGAATGAGTATTTATCTGTCGAGTCTCTG-ATCTCTCTCTTTAAATGAGACA 109558
 QY 7426 GTAAATCCCATGAGGAGTGTGGGAGAAATCAGAGATCAAAACGCTGTATACAT 7485
 DB 109557 GTAAATCCCATGAGGAGTGTGGGAGAAATCAGAGATCAAAACGCTGTATACAT 109498
 QY 7486 CTGCTTCTGTTTCCAGGCTCATGAGCTGGGTTTCTGAGCAATGATTCACATCCCA 7545
 DB 109497 CTGCTTCTGTTTCCAGGCTCATGAGCTGGGTTTCTGAGCAATGATTCACATCCCA 109438
 QY 7546 GCTTGGGTACAGAACTGACACCAATCAAGAGAGTGTGAGAGACTCTTGTCTTACAGAG 7605
 DB 109437 GCTTGGGTACAGAACTGACACCAATCAAGAGAGTGTGAGAGACTCTTGTCTTACAGAG 109378
 QY 7606 ACCCTGAGCTTACAGGAGCTGACGTCATGCTTCCCTTGTGCGCTGACAGAAACGG 7665
 DB 109377 ACCCTGAGCTTACAGGAGCTGACGTCATGCTTCCCTTGTGCGCTGACAGAAACGG 109318
 QY 7666 GTTGTGCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7725
 DB 109317 GTTGTGCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 109258
 QY 7726 AACCTGGTGGGCGGAGCTTCTCTTCTTAGCGGCGCAATATATGTTGCGCGTTACGC 7785
 DB 109257 AACCTGGTGGGCGGAGCTTCTCTTCTTAGCGGCGCAATATATGTTGCGCGTTACGC 109198
 QY 7786 CTCAATCAATATCCGCAATCCCATCTCCAAATCTCTGAGTCTGATGATCCTTCCCTAC 7845
 DB 109197 CTCAATCAATATCCGCAATCCCATCTCCAAATCTCTGAGTCTGATGATCCTTCCCTAC 109138
 QY 7846 TTTTATAGGCTTAAGCACTGAGGCGCATGAGGCGCATGAGGCGCGCTGCTGCTGCTGCTG 7905
 DB 109137 TTTTATAGGCTTAAGCACTGAGGCGCATGAGGCGCATGAGGCGCGCTGCTGCTGCTGCTG 109078
 QY 7906 CCATCTGATGACACTGCGC--CGCCCAAGATCTGTCATCGGTATGATGCTGCTGCTGCTG 7963
 DB 109077 CCATCTGATGACACTGCGC--CGCCCAAGATCTGTCATCGGTATGATGCTGCTGCTGCTG 109018
 QY 7964 TCTGGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8023
 DB 109017 TCTGGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108958
 QY 8024 GTGATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8083
 DB 108957 GTGATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108898
 QY 8084 TTTTATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8143
 DB 108897 TTTTATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108838
 QY 8144 CCGGAAAGATGCGGCTGACCAAGCTGTGACATCTCTTCAACAGTGTGCTTCTC 8203
 DB 108837 CCGGAAAGATGCGGCTGACCAAGCTGTGACATCTCTTCAACAGTGTGCTTCTC 108778
 QY 8204 TCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8263
 DB 108777 TCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108718
 QY 8264 GGAAGCTTATTTTGTGATGTCATGATGATGATTTTCTGCTGCTGCTTCAACAGCA 8323
 DB 108717 TGAAGTCAATTTTGTGATGTCATGATGATGATTTTCTGCTGCTGCTTCAACAGCA 108658
 QY 8324 GTGCAACCCCATCATTTACTTCTTGTGAGCTTCTTTAGCGAGCTCAAAATGAGACA 8383
 DB 108657 GTGCAACCCCATCATTTACTTCTTGTGAGCTTCTTTAGCGAGCTCAAAATGAGACA 108598
 QY 8384 ACCTGAAGCTGCTTCAAGGCTCTGAGACACGCTGAGGCTGATGAAAGTGTGAG 8443

Db 108597 AACTGAGCTGGTCTCCAGAGGCGCTGCGAGGACATGCTGAGG-----TGAAGGTGAG 1085482

QY 8444 GGTGGCTTCTCAGAGAACTTGGAGCTGTGGGAAAGCAGATTGGAGCAGTGAAGAA 8503

Db 108541 GGGCGCTTCTGAGGAAACCTGAGAGCTCGGGAAGCAGATTCCGGCAGTGAAGAA 108482

QY 8504 CCTCGGCTCTCTGAGAGGAGCTTGGAGGCAATGCTGCGCCAGCAGCTTGAACA 8559

Db 108481 CCTGCGCTTCTGAGTCAAGCAGGAGCTTGGAGGCAATGCTGCGCCAGCAGCTTGAACA 108422

QY 8560 TTATATGATTTTCTTACCTTCTGCTCAGAAATGCTCAGAG 8604

Db 108421 TTACATGC-GTTTCTTACTGTTGTGCTCAGAAATGCTCAGCG 108378

RESULT 5
AC139482 165490 bp DNA linear HTG 04-FEB-2003
LOCUS Homo sapiens chromosome 5 clone RP11-583F24, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.
ACCESSION AC139482.1 GI:28201512
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 165490)
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 165490)
DOE Joint Genome Institute.
REFERENCE Direct Submission
TITLE Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1518520
Center clone name: RPCI-11_583F24

Summary Statistics
Consensus quality: 147035 bases at least Q40
Consensus quality: 155210 bases at least Q30
Consensus quality: 158567 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 164390; sum-of-coverage estimation
Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
Quality coverage: 4.19 in Q20 bases; sum-of-coverage estimation
NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1249: contig of 1249 bp in length
* 1250 1349: gap of unknown length
* 1350 3680: contig of 2331 bp in length
* 3681 3780: gap of unknown length
* 3781 6668: contig of 2888 bp in length
* 6669 6768: gap of unknown length
* 6769 10866: contig of 4098 bp in length
* 10867 10966: gap of unknown length
* 10967 17019: contig of 6053 bp in length
* 17020 17119: gap of unknown length
* 17120 22433: contig of 5314 bp in length
* 22434 22533: gap of unknown length

22534 34645: contig of 12112 bp in length
* 34646 34745: gap of unknown length
* 34746 49173: contig of 14428 bp in length
* 49174 49273: gap of unknown length
* 49274 67058: contig of 17785 bp in length
* 67059 67159: gap of unknown length
* 67159 84928: contig of 17770 bp in length
* 84929 85029: gap of unknown length
* 85029 114464: contig of 29436 bp in length
* 114465 114564: gap of unknown length
* 114565 165490: contig of 50926 bp in length.
* Location/Qualifiers
1. 165490
source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-583F24"
/clone_11b="RPCI human BAC library 11"

ORIGIN
Query Match 62.4%; Score 5380.4; DB 2; Length 165490;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 6838; Conservative 0; Mismatches 906; Indels 394; Gaps 52;
Query 763 AAAATCTTAAGAAATTTCTTACCCAGGAGGTGAAGGCTCTCAAGAGAACTACAA 822
Db 114643 AAACCACTTAGAGGTTTAGCAAAAAGAGGTGAAGGCTCTCAAGAGAACTAT 114698
QY 823 AACACAGCTGACATCATAGATGACACAAAGTGAAGAACATCCATGCTCATGATG 882
Db 114699 GAAAGCTGAATCATATGATGACACAAAGTGAAGAACATCCATGCTCATGATG 114758
QY 883 GGTGAATCATATTTGTGAAGTGAACATTTGCTCCAAAGCAATTTACAGTTCAATGA 942
Db 114759 GGTAAATCATATTTGTGAAGTGAACATTTGCTCCAAAGCAATTTACAGTTCA 114818
QY 943 ATTCCACCAAAATATCATCATCTTCTTACAGAACTGAGAAACAAATCTTAAAT 1002
Db 114819 ATTCCACCAAAATATCATCATCTTCTTACAGAACTGAGAAACAAATCTTAAAT 114877
QY 1003 CATATGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 1062
Db 114878 AAT-----ACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 114922
QY 1063 AAAAGACAAATCTGAGGACATCATTAATCCATCTTCAATCTATCTACAGCTATA 1122
Db 114923 AAAAGACAAATCTGAGGACATCATTAATCCATCTTCAATCTATCTACAGCTATA 114982
QY 1123 ATACCAAAACATCATGACATGACATTAATCTGAGGACATCATTAATCTGAGGACAT 1182
Db 114983 GTACCAAAACATCATGACATGACATTAATCTGAGGACATCATTAATCTGAGGACAT 115042
QY 1183 GAGAAATCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1242
Db 115043 GAGAAATCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 115102
QY 1243 AATATTAATGAGGAGAAACATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1302
Db 115103 AATATTAATGAGGAGAAACATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 115162
QY 1303 ACATGTGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
Db 115163 ACATGTGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115221
QY 1363 ATGATCAAGAACTTAATCTGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1422
Db 115222 ATGATCAAGAACTTAATCTGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 115281
QY 1423 AAAAATGCTTCTGAGATTCATTAAGCAAAAGCTTCAATGAGCAAGAAACCCAAAG--TA 1480
Db 115282 AAAAATGCTTCTGAGATTCATTAAGCAAAAGCTTCAATGAGCAAGAAACCCAAAGGATTA 115341

OY	1481	AATGCAACAAAAACAAAATTAATAGATGAGCTTAATTAATTAAGAAAGCTTTGGCA	1540
Db	115342	AATGCAACAAAACAAAATTAATAGATGAGCTTAATTAATTAAGAAAGCTTTGGCA	115401
OY	1541	GCAAAAACATCATTTAGCAGACAAACAGACAACCCAGAGTGAAGAAATCTTCACA	1600
Db	115402	GCAAAAGAAATCATCGACAGAAACAAAGAGTAATCCACGTAGTGGAGATTAATCTTCACA	115461
OY	1601	AACTAAGCATCTGACTAAGACTAATATCCGAATTCACAGGAATCTCAACAAATCAGC	1660
Db	115462	AACTATGCACTCGACAAAGAGCAAAATATCAGATCTTACAGGAATCTCAACAAATTAGC	115521
OY	1661	AAGAAGAAAGCAAAACAATCCATGAAGAAGTGGGCTTAAGGACATGAAATGACAATCTCA	1720
Db	115522	AAGAAGA-----AAACGTACATCAAAAGTGGGCTTAAGGACATGAAATGACAATCTCA	115577
OY	1721	AAAGAAGATATACAAATGCGCAACAAC--AGGAAAAATGCTTAAATCATCATATGATTT	1778
Db	115578	AAAGAAGATATACAAATGCGCAACAACAATGAAAAATGCTTAAATCATCATATGATTT	115633
OY	1779	AGGGAATGTAAATCAACCTGTATGCGATACACCTTACTCTGCAAGATGTGCATA	1838
Db	115638	AGGGAACACCAATCAACAAACCAATTAATGCAATGCACTTACACCTGCGAAGATGCGCAT	115697
OY	1839	ATTTAAAAATCTAAAAATATATATGATTTGTTGGTGTCTGTGCTGTATTAAGAACATTTTA	1898
Db	115698	ATTTAAAAATCTAAAAATATATATGATTTGTTGGTGTGTGTGCTGTATTAAGAACATTTTA	115757
OY	1899	CAGCTCTGGTGGGAATGTAAACTTGGCGCAACCACTATGGAACAAAGTGTGAATTTCTT	1958
Db	115758	CAGCTCTGGTGGGAATGTAAACTTGAACAACAACCACTATGGAACAAAGTGTGAATTTCTT	115817
OY	1959	AAGGAATTAAGATGATGACCAATTTGATTCAGCAATCCCA-----	2000
Db	115818	AAAGAACTAGAGTATGATCTCAAACTTGTATCGAGTAATCCCACTCTGGGTATCTACCA	115877
OY	2001	-----	2000
Db	115878	GAGAAAAAGAGTCAITTTTACGAAAAAGACACTTGACCTTGATGTTTATGACAGACAA	115937
OY	2001	-----	2000
Db	115938	TTGCAATTAACAAAATATGGAACACAGCTCAAAATGCCATCAATCACCAAGTGATTAAGA	115997
OY	2001	-----	2000
Db	115998	AAATTAATATATATATATGTTGTGTATGTATATGTATATATATATATATATATATATAT	116057
OY	2001	TTAAATATGTATTAATATATATATTT-----ATATACATGGAATACAA	2044
Db	116058	GTATATATATTAAT	116117
OY	2045	CTCAGCCATTAATAAAGAAATTAATAATGATTCATTCAGAGCAATGTAGATGGAATTTGAGAC	2104
Db	116118	CTCAGTAATTAATAAAGAAATGAAATTAATATATATATATATATATATATATATATATAT	116177
OY	2105	CTTATTTCTAAGTGGGTAATCTAGGAATGGAACCAACCAATCATATATGTTCTCACCTTAC	2164
Db	116178	CATTATTTCTAAGTGAAGTAACTCAGGAATGGAACCAACCAATCATATATGTTCTCACCTTAT	116237
OY	2165	AAAGTGGGGCTTAAGCTGTGAGACACAGAAAGCATAGAAATGATTAATGAACCTGGGGAC	2224
Db	116238	AAAGTGGGGCTTAAGCTGTGAGACACAAAGGCTAGAAATTAATCAATTAATGAACCTTTGAGAC	116297
OY	2225	TTGAGGGGGAAGATGGAAGAGAGGCGAGGATTAAGAATCAACATATGGTGAACGTGTAC	2284
Db	116298	TTGAGGGGGAAGATGGAAGAGAGATGAAGATTAAGAATCAACATATGGTGAACGTGTATAC	116357
OY	2285	ACTGCTCAGGTATGAGGTCAACCAAAATCTCAGAAATTAACCTTAAGAAGCTTATCCATG	2344
Db	116358	ACTGCTTGGGTATATGAGGTCAACAAATCTCAGAAATTAACCTTAAGAAGCTTATATCATG	116417
OY	2345	GAAGCAACACCACTGTTCCCAAAATCCCAATGAATTAATAATTAATTAATGA	2404

Db	116418	TAACCAACCACTGTTCCCAAAAACCATGGAATGAC--AATATATATAATGA	116474
-	2405	TTTAATTTACAGAA-TTTAAAAAGTTCACTGTTCAAGATTATATATATGAAATGA	2463
OY	116475	TTTGATTTACAGAACTTAAAAAAGTTAGCTGTTCAAGTTGATGA--AAAGAAATGAAA	116532
Db	2464	ATGAAATATGTAAGCAAGTGTAGCTCTGAGCAATGGGACCTAGATTTCACCTTGCAAT	2523
OY	2524	ACACTTCTCTGGCACTTTGAAAAAAGATACATCATGTAATATATACCACTATGATAAG	2583
Db	116533	ATGAACAGTGTAGCAAGTGTAGCTCTGAGCAATGGGATTCGACTTTCACCTTTCAT	116592
OY	2584	AAAAATC-AAAAAATGTGTAGAGCCATTGTAGCCTTGAATGTGTCCATGATCTACT	2641
Db	116653	AAAAACAAAAAATGTGTGTCAGGACATGTGTAGCCTTGAATGTCTATGATCTACT	116712
OY	2642	TTTTCAATTTGGATATAAAGCCTCATATGATGTTCACATTCCTTAATGTGATGCTTAG	2701
Db	116713	TTTTTCATTAATTAATAAAGCCTCATATGATGTTCACATTCCTTAATGTGATGCTTAG	116772
OY	2702	CCCATTAATGATTTTAAATCAGACAGCAATTACTTAACAGAAAGTTGAACAAGATGG	2761
Db	116773	CCCATTAATGATTTTAAATCAGACAGCAATTACTTAACAGAAAGTTGAACAAGATGG	116832
OY	2762	ACGTGATAGAGAGGCTTAAATGTACTGATATGGAACAGAGCCAAAGATCATCTCAGT	2821
Db	116833	ACGTGATAGAGAGGCTTAAATGTACTGATATGGAACAGAGCCAAAGATCATCTCAGT	116892
OY	2822	TAGATTTGTGTCTCAAAATACCTCTGGCCTCTGATTTGCCATPATGCTCTATACAGAAA	2881
Db	116893	TGGATTTGTGTCTCGAACTTCTCTGGCCTCTGATTTGCCACAGTCTCATGACAGAAA	116952
OY	2882	TAAACAAGCTGTCCAGCATCTTCTGTAAGCCTGGAATGGCTCACAGCTTCATTTAGCTC	2941
Db	116953	TAAACAAGCTGTCCAGCATCTTCTGTAAGCCTGGAATGGCTCACAGTTCATTTAGCTC	117012
OY	2942	CTGTAGCATCTCTCGAATTAAGCAACAGAAAAGTCTCTGAAGTCACTGAATCCAG	3001
Db	117013	CGTGTAGCATCTCTCGAATTAAGCAACAGAAAAGTCTCTGAAGTCACTGAATCCAG	117072
OY	3002	AAAGGCTCTTACCTTTAGCAACAAGAGGTCTTCAACACTGAGCA---AAGAGGAAC	3057
Db	117073	AAAGGCTCTTACCTTTAGCAACAAGAGGTCTTCAACACTGAGCA---AAGAGGAAC	117132
OY	3058	GATAAGGTAATGTAACCAAGAACTCTCTTCCACAAGTCAATTAATGATTTTGTGCTAAG	3117
Db	117133	GACAAGGTAATGTAACCAAGAACTCTCTTCCACAAGTCAATTAATGATTTTGTGCTAAG	117192
OY	3118	ATCATGTCTTATGTCTTCCACCTTGATGCTATACAGAGGGGCTCAAGACTTGTTCAG	3177
Db	117193	ATCATGTCTTATGTCTTCCACCTTGATGCTATACAGAGGGGCTCAAGACTTGTTCAG	117252
OY	3178	AAAAAGACAGACATGAAGCTTCTTTCAGAACTGAGTGTGTCAACCAACTGTGT	3237
Db	117253	AAAAAGACAGACATGAAGCTTCTTTCAGAACTGAGTGTGTCAACCAACTGTGT	117312
OY	3238	GAGCTTAAATGTGTCCCCCTTTCTAATTTATCTTCCCATATCACCTCTTCATTCCAA	3297
Db	117313	GAGCTTAAATGTGTCCCCCTTTCTAATTTATCTTCCCATATCACCTCTTCATTCCAA	117372
OY	3298	TCATTCATTCGCCCTCATGAGAGAAATGCTACTTAACATTTCAATTAATGACAGAG	3357
Db	117373	TCATTCATTCGCCCTCATGAGAGAAATGCTACTTAACATTTCAATTAATGACAGAG	117432
OY	3358	GACATGACAGGATTTCTTCCAGAGTTGAATGCTATAG-AAGCAATTTCTTGTGTTTCA	3416
Db	117433	GACATGACAGGCTTCTTCCAGAGTTGAATGCTATAG-AAGCAATTTCTTGTGTTTTC	117492
OY	3417	TTACTTTCAAAATTTATTTCTTGTGCTATATGTGAAAGGTCTTAAGAAATATATAGTGG	3476

Db	117433	TTACTTTTCCAAATTTTCTCTCTCCCTGCTATCTGGAAAAGTTTCAAGGAACATGGAATGG	117552
Qy	3477	CCC-----AATTAATTAAGAGCTGTTTTCATGAGAAAGTAATTTACAAAGATGCACAGAG	3529
Db	117553	TTCCAAACTGAAATGATGTAGAGAGTGTTTCTCATGAGGAGAAAGTATTTACAAAGACATGACAGGG	117612
Qy	3530	TTAA-----GGGTACAGATCTCTAAGACAGCAATATACATAGGGGAGAGCATTA	3572
Db	117613	TTTAAGAAAAGAAAAATGACATGGGTGACGCCCTTAAGACAGCAATATCTAGGGGAGACATA	117672
Qy	3573	CTTCCCTCCCTTAGAGCTGAAACCGACACAGGAGAGAGAGATTCACATTTGTC-----GCCAT	3626
Db	117673	AGTCTCTCCCTGACGTGAAAGGAGACAGGAGAGAGAGATTAATTTATGTGCACCATACCAT	117732
Qy	3627	AGCCATAGCTGTAGCCATTAAGGCTGGAGAGCATGACAGCGCAAGTGGAGAACCTTCGG	3686
Db	117733	AGCCATAGCTGTATCATATAGGGGTGGAGAAACAAGAGCGGACAGGTGGAGAACCTCTGTA	117792
Qy	3687	TGGCCAAACCCACAGCCACACAGGCTGATATATAGTTTGATCTGTGTTCCACCAAAATCTC	3746
Db	117793	TGGCCAAACACACAGCCACACAGGCTGATATATAGTTTGATCTGTGTTCCACCAAAATCTC	117852
Qy	3747	AGTTGATATTGTAATTTCCATGTGTGGAGAAAGGGCCCTGTGGGAGATGATTTAATGATCA	3806
Db	117853	ATGTTGATTTGTAATTTCCCATGTGTGGAGAAAGGGCCCTGTGGGAGATGATTTAATGATCA	117912
Qy	3807	CGGGAGTGGTTTGTGATGATGATTTTAAACCATCCCTTTGGTATGTGTGTGTGATTA	3866
Db	117913	CGGGAGTGTTTTCATATGATGATGTTTAAACCATCCCTTTGGTATGTGTGTGTGATCA	117972
Qy	3867	CTGACGAGTGTTCATGAAATCTAAGTTGTTTAAAGGTGTAGACCTCTCTCC--CTCTCT	3924
Db	117973	CTGATGAGTGTTCATATATATCTAAGTTGTTTAAAGGTGTAGACCTCTCTCTCTCT	118032
Qy	3925	CTTATCTCCCTGCTCTACATGTGAGAAGCCCTCGCTCCCTTTGGCTTTCAACAGGATTTG	3984
Db	118033	CTTGTCTCTGCTTTTCCATGTGAGATGCTCTCTCCCTTTGGCTTTTCAACATGAAATG	118092
Qy	3985	GAAGCTTCTGTAGGCTCTCCACAGAACAGAGCTGTATGCTTCTGTACATCTGTAGA	4044
Db	118093	GAAGCTCTGTAGGCTCTCTTAAACAGAAAGCTGTATGCTTCTGTATATAGCTGTAGA	118152
Qy	4045	GCTATTAAGCAGTTAAACCATTTCTCTCTAATAATTTCCAGTCTCAGGTAATTTCTTTT	4104
Db	118153	GCCATGAAACCAATTAACCTGTTTCTTCTAATAATTTCCAGTCTCAGGTAATTTCTTTG	118212
Qy	4105	AGCAATTTGAGATGAACTAATACACAGACAGAGAGCAGAGATGGAATTTCCAAAGTGT	4164
Db	118213	AGCAATTTCAAGATGAACTAATACCTGACAGAGAGCCAGAGAGTGGAAATTTCCAAAGTGT	118272
Qy	4165	CTTTCCTGTGTCCTTCCAGTCTCTGTGTGTCTCCAGTGTCTCAATTTCCACAGAAA	4224
Db	118273	CTTCTCTGTGTCCTTCCAGTCTCTCCCTGTGTGTCTCCAGTGTCTCAATTTCCACAGAAA	118332
Qy	4225	CCAGAAATTAAGAAATCCCATCTGATGTGTACATAGAACCACTCTCTTGGGATGTCAA	4284
Db	118333	CAAGATTTTAAGAAATCCCATCTGATGTGTACATAGAACCACTCTCTTGGGATGTCA-A	118391
Qy	4285	ACAGGATTAAGAAAGATGGAACCAATCTCTCAGTGAATGAGCATATCCCTCAACCT	4344
Db	118392	ACAGGATTAAGAAAGATGGAACCAATCTCTCAGTGAATGAGCATATCCCTCAACCT	118451
Qy	4345	TCTTGTATCTCTTAATCTCTGTGGGGCTTCTCTATC-----TGATGTATCCCTGTCTCA	4398
Db	118452	TCTTATATCTCTTACTCTCTGTGGTTTTTCTGTCTTAAAGGTGATTTGATCTGTGTCTCA	118511
Qy	4399	TTTCACTCTATCAGACTACTTTAATGTTTGGCTTGTCTTCTCTACGTCACTTTTATG	4458
Db	118512	TTTCACTCTATCAGAC-----TAATGTTGGATGTCTTCTTACTGTCACTTCTATG	118566
Qy	4459	CAGAAATGTTTGATTTGTTTAAATGCAATGAAATTAATTAATTTTAAAGAA--4515	
Db	118567	CAGACATGTTTGGTTCATCTCAAAATGCGTAGAAATTAATATGTAATTTTAAAGGAG	118626

QY	4516	ACATATGATTTTTGTTTGAATATTAATTAATTTGGCTGATCTTAATAAGACATGAAGAAGAA	4574
Db	118627	GACATATTTCTTTGTTTGAATATTAAGTTGGTTGGTCTCTCAATTAAGACAGAACTGAA	118686
QY	4575	ATATCTTAACAGAAAGATATAGTTGTCCTCTGGGTCACTAGTTCGAACTCAAGAT	4634
QY	4635	TCAACAACTACAGAGGAACTTTCCAAAAATTAAGTGTGGGGAGTTGTATATGA	4694
Db	118746	TCACAAACCCACAGGTGAAACTATTTAAAAATTAAGTATGGGGAGTTGTATATA	118805
QY	4695	CTGAACAGGTCAAACTGTATTTCTTTGTCACTATTTTCTGAANAACCTACAAT-ATAACA	4753
Db	118806	CTGACAGATATAACTGTGTTTTCTTGCTCATTACTTCCTGAATTGATAGATAAAAAAA	118865
QY	4754	AGAACTATATAGCACTTTGGATTTTGTCACTTATCTTAATACCTT-TAATGATTTAAT	4812
Db	118866	AGAAATTACATAGCAATTAACATTTTGTCACTTATCTTAAGTAACTAGAAATTAATTAAT	118925
QY	4813	GTATCTGGAGAAAGATGCAATAGATATATATACAAATACATAT-----ATAGGAAAT	4865
Db	118926	GTATCTAGAGAAAGATGCAATAGGTGTATACAAATATCACTCGTTTGTATTAAGAAAT	118985
QY	4866	GAGCATCTGCAGATTTTGGTCTGTGTCCTGGGTTCTGAAAGAAATCCCTGTAAATACACA	4925
Db	118986	GAGCATCTGCAGATTTTGGTGTGTCTGTAGGTTGGGAAAGCAAGCCCCAATGATACCCA	119045
QY	4926	AAATATGACACTTTTGGAGATCTGAACCTAGAAAGCTCCAAAGCATCATCATCGAATTTCCA	4985
Db	119046	GAGATGACCACTTTCTTAATCTCCAGACTGGAAAGCTCCAAAGCACTAATTCAGAAATTC-	119104
QY	4986	AAATATGCTGCTCCCAAGTCTCTAGAGAGTTGCCCTCATCTTGATCTCATAGTTC	5045
Db	119105	AAATATGAAACCTCCCAAGTCTCTAGAGAGTTGCCCTCATTTTGTGATCTCTGATGATTC	119164
QY	5046	CCAGGCACTTAGCACTTCCAGTCTTATGGAAAAAGACGAGGGGAAAGAGAGCTTTGCT	5105
Db	119165	CCAGGCACTCGGCATCTTCAGCATATGG-AAAAGATGAAAGAAAGAGAGAGCCCTTGCT	119223
QY	5106	CCCTTATTAATCCCATAGAGCCAGAGACTTGCTGTCACTTTGATCTTCCACTTA	5165
Db	119224	TGTTCTATTTGATCCCATAGGCCAGGGA-----TACTTTTGTCACTTACATTTGA	119274
QY	5166	ACAGCACTGCTCATGGAGTGTCACTCAAGCATCAAGAAAGCTGGATGGGTCTTGCT	5225
Db	119275	AGAGTCTCGTCTCATATGATGCA-CAATCATCAAGAAACACTGGGATGTGGGTCTTGCT	119333
QY	5226	GCTGCTTGTACATTTCTCAAAAAGTTATCTGACCAAAAAGAAAT-----CTGGGGGGA	5280
Db	119334	GCTGCTTTTACCTCTCAAGACTGTATATGTGCAAAAGGAAATGATTATTTGGGGCAA	119393
QY	5281	ACCAACAGTCTTTCAGCCCTCGACTGTCTCGATTTCTGTGCTCACTCAAGATTTTTCA	5340
Db	119394	CCTACAGTCTTTCAGCCCGGCACTGTCTCGATTTCTGTGCTCACTCAAGATTTTTCA	119453
QY	5341	GGAACTCTCAGAAATTAATTAATGTGGGGCAGAGAACAACTGGAGTCTGTCGAGA	5400
Db	119454	AGAACTCTCAGAAATTAATTAATGTGGGACAGAGAAACAGTTCAAGAAAGTCTCAACAGGG	119513
QY	5401	CTCCAGGGGACCAAGGGGCTGTATTTGAACTGCTCTTCAATGTTGTGTAAC-CAGAGAAAGCC	5459
Db	119514	CTCCAGGGGACCAAGGGGCTGTGTATTTGAACTGCTCTTCAATGTTATGATCTCAGAGAAAGCC	119573
QY	5460	TTTAATTTCTAGGCTTAGCTTCACTTATGTTATATAGAGATATACCATAGACGTC	5519
Db	119574	TTTAATTTCTATAGGCTTAGGCTTCACTTATGTTATATAGAGATATACCATATAGATGTC	119633
QY	5520	TTTAAGAAGCATATAGCATGTTTAAACAACATGCTTAATGTTGGTGTATCACACATGAAA	5579
Db	119634	TTTA-----GAAAGCATGTAAATGTTGATGATATCTACATGAGAAA	119675

QY 5580 AAGCAGGACGACTTACTCTTACGATCTTGGGTTTCATGAGAACAACTATC 5639
 DB 119676 AAGCAG-----CTATGAGCTTCAGGTTTCATGAGAAATATATATC 119722
 QY 5640 ATACCATCACTATGATGAGCAAAACAGTTAGTCTGATGATGATATACAGAGTTCT 5699
 DB 119723 ATACCCATATATGATGAGCAAAAGTTTAGTCCCTGATATGATATACAGATGTTCT 119782
 QY 5700 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGGCAGGTTCTCCCTAGA 5759
 DB 119783 CTTTCTCTCCATTTCTTTGGGCAATCAGATATATGAGAGCTGTCTTCTACTA 119842
 QY 5760 GAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5819
 DB 119843 AAGGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119902
 QY 5820 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGGCAGGTTCTCCCTAGA 5879
 DB 119903 CTTTCTCTCCATTTCTTTGGGCAATCAGATATATGAGAGCTGTCTTCTACTA 119962
 QY 5880 GCCTCTGATGAGCAAAAGAACTCTGATTTTCTGTAATGAGCAGCCTGCTCT 5939
 DB 119963 GCCTCAGATGAGCAAAAGAACTCTGATTTTCTGTAATGAGCAGCCTGCTCT 120022
 QY 5940 AT-TTTCCTTATTTGATTTGCTTCTCTATCTGATGATGATGATGATGATGAT 5998
 DB 120023 ATATGTTCTCTATTTGATTTGCTTCTCTATCTGATGATGATGATGATGATGAT 120082
 QY 5999 TTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGGCAGGTTCTCCCTAGA 6058
 DB 120083 TTTTCTCTCCATTTCTTTTGGGCAATCAGATATATGAGAGCTGTCTTCTACTA 120142
 QY 6059 CTTTACCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6118
 DB 120143 CTTTACCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120201
 QY 6119 GAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6178
 DB 120202 GAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120254
 QY 6179 CCCCCGACAAATATATCTCTCTTATATCTTATATCTTATATCTTATATCTTAT 6238
 DB 120255 TCCGAGGACAAATATATCTCTCTTATATCTTATATCTTATATCTTATATCTTAT 120313
 QY 6239 AATATATATTTTATCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 6298
 DB 120314 AATATATATTTTATCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 120373
 QY 6299 TGTCAATCAGCTGCTGATTTATCTTCCCAATTAATGATGATGATGATGATGATGAT 6358
 DB 120374 TGTCAATCAGCTGCTGATTTATCTTCCCAATTAATGATGATGATGATGATGATGAT 120433
 QY 6359 TTGCATGATATCTCTCTTATATCTTATATCTTATATCTTATATCTTATATCTTAT 6418
 DB 120434 TTGCATGATATCTCTCTTATATCTTATATCTTATATCTTATATCTTATATCTTAT 120492
 QY 6419 CTTCTCTCATACAAATGTTGATCAACAAAGAACTGATCAAAAGATCTTCCGAAAGAG 6478
 DB 120493 CTTCTCTCATACAAATGTTGATCAACAAAGAACTGATCAAAAGATCTTCCGAAAGAG 120551
 QY 6479 AGAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6538
 DB 120552 AGAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120611
 QY 6539 ATTCCACATTCATATACCTGCTGATGATGATGATGATGATGATGATGATGATGAT 6598
 DB 120612 ATTCCACATTCATATACCTGCTGATGATGATGATGATGATGATGATGATGATGAT 120671
 QY 6599 CTTCTCTCATACAAATGTTGATCAACAAAGAACTGATCAAAAGATCTTCCGAAAGAG 6657
 DB 120672 CTTCTCTCATACAAATGTTGATCAACAAAGAACTGATCAAAAGATCTTCCGAAAGAG 120731
 QY 6658 AAAATGCTTCTGAGGCTTCCGCAAGCAGCAGTATGCTTATTTTAAAGCAGATCAGT 6717

DB 120732 AAAATGCTTCTGAGGCTTCCGCAAGCTGATCAGATCTGATGATGATGATGATGAT 120791
 QY 6718 GGTCAATTTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACAGATGATGATGATGAT 6777
 DB 120792 GGTCAATTTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACAGATGATGATGATGAT 120851
 QY 6778 ACAAATGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6837
 DB 120852 ACAAATGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120911
 QY 6838 CTTTACATGCTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 6897
 DB 120912 CTTTACATGCTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 120971
 QY 6898 CTTTACATGCTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 6957
 DB 120972 CTTTACATGCTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 121031
 QY 6958 ATGG-ATGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7016
 DB 121032 ATGGAATGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121091
 QY 7017 CACTGATTTATTTTCTCAACCCAGCAGGCTCCGCAATTTTCAAC--ATCTAGCGAAGATC 7074
 DB 121092 CACTGATTTATTTTCTCAACCCAGCAGGCTCCGCAATTTTCAACATTTATCTGACAAAGGTC 121151
 QY 7075 CCAATTTCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7134
 DB 121152 TCAATTTCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121211
 QY 7135 GTCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7194
 DB 121212 GTCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121270
 QY 7195 TTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7254
 DB 121271 TTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121330
 QY 7255 TTTGATCTTATGTTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 7314
 DB 121331 TTTGATCTTATGTTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 121390
 QY 7315 GGTTCAGAGTCAACAAAGCTGATTTCAACCTGATTTGAGGACCCCACTTTTGTAT 7374
 DB 121391 GGTTCAGAGTCAACAAAGCTGATTTCAACCTGATTTGAGGACCCCACTTTTGTAT 121449
 QY 7375 AGTGATCTTATGTTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 7432
 DB 121450 AGTGATCTTATGTTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 121509
 QY 7433 CCAATGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7492
 DB 121510 CCAATGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121569
 QY 7493 CTTGTTCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7552
 DB 121570 CTTGTTCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121629
 QY 7553 GTACAGAACTGACCAATCAACAGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 7612
 DB 121630 GTACAGAACTGACCAATCAACAGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 121689
 QY 7613 GCTTCAAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7672
 DB 121690 GCTTCAAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121749
 QY 7673 TTTGCTCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7732
 DB 121750 TTTGCTCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121809
 QY 7733 TCCGAGGCACTTCTTCTTATGAGGCAATATATGATGATGATGATGATGATGATGATGAT 7792

D	b		121810	TGAGGGCAGACTTCCTCCCTTCCACAGGGCGCACGTATTAACATTCCGGCTCACTCCATCA	121869
O	y		7793	ATATCCGCATCTCCATCTTCGAATAATCTCATGTCCTGTATGACTTTTCCCTACTTTATAG	7852
D	b		121870	ATATCTGCATCTCCATCTTCGAATAATCTCATGTCCTGTATGACTTTTCCCTACTTTACAG	121929
O	y		7853	GCCCTAAGATGCTGAGCGGCATCAGACAACGAGCGCGCTGCCCTGCATCCGTGGGCCATCT	7912
D	b		121930	GCCTAGCTTCTTGATGTCAGATGAGACCGAGCGCTGCTTGTCGTCCTGGCCCATCT	121989
O	y		7913	GGTACCACATGCGCG--CGCCCCAGATACCTGTCTATGATGATGATGATGATGATGATGATG	7970
D	b		121990	GGTACCGCTGCTCTCTCCCCACACACCTGTGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTG	122049
O	y		7971	CCTGTCCCTGCTGCGGAGTATCTTGAGTGATGTTCTGTGACTTCTCTGTTTATGTGTGC	8030
D	b		122050	CCTGTCCCTACGCGGAGCATCTTGAGTGAATGTTCTGTGACTTCTCTGTTTATGTATGC	122109
O	y		8031	TGATCTGTGTGTGTGAACGTCATATTTCAATCAATGAGCGGTGTGTGTGTGTGTGTGTGTGTGTG	8090
D	b		122110	TGATCTATTTGTGTGTGAACGTCATATTTCAATCAATGAGCGGTGTGTGTGTGTGTGTGTGTGTGTG	122168
O	y		8091	TGTGTGTCTCTG	8150
D	b		122170	TGTGTGTCTCTG	122229
O	y		8151	GATGCGCTGACCAAGGCTGTACGTACATCTCTCTCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	8210
D	b		122230	GATGCGCTGACCAAGGCTGTACGTACATCTCTCTCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	122288
O	y		8211	TGGCGTGGCTTG	8270
D	b		122290	CAGCTGCGCTTG	122349
O	y		8271	CTTATTTG	8330
D	b		122350	CATTTCCTG	122409
O	y		8331	CCCACATATTATCTTCTG	8390
D	b		122410	CCCCGATATTATCTTCTG	122469
O	y		8391	GCTGTGTCTTCCAGAGGCTCTGTGAGACACGCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	8450
D	b		122470	GCTGTGTCTTCCAGAGGCTCTGTGAGACACGCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	122529
O	y		8451	TCCTCAGAGAAACCTGTAGAGCTG	8510
D	b		122536	TCCTCAGAGAAACCTGTAGAGCTG	122588
O	y		8511	CCT----GTGACAGAGACTTTGAGAGCAATGTCTGCCCTGCACACCTTGTACAAATTATATG	8566
D	b		122586	CCTGTACAGTCAAGCGGAGCTTTGAGAGCAATGTCTGCCCTGCACACCTTGTACAAATTATATG	122645
O	y		8567	CATTTCCTTGTAGCTTCTG	8604
D	b		122646	C-GTTTTCTTG	122682
RESULT 6					
A	C	AC020568/c			
L	O	AC020568	192618 bp	DNA	linear HTG 17-AUG-2000
D	E	Homo sapiens chromosome 20 clone RP11-206Cl, WORKING DRAFT			
A	C	AC020568			
K	E	AC020568.4 GI:9280789			
S	R	HTG; HTGS_PHASE1; HTGS_DRAFT.			
O	R	Homo sapiens (human)			
O	R	Homo sapiens			
O	R	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;			
O	R	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
R	E	1 (baaes 1 to 192618)			
A	A	Waterston,R.H.			

```

JOURNAL TITLE
REFERENCE Unpublished
AUTHORS The sequence of Homo sapiens clone
TITLE 2 (bases 1 to 192618)
JOURNAL Waterston, R.H.
Submitted (04-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT ON Jul 19, 2000 this sequence version replaced gi:2211362.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0206C01
----- Summary Statistics -----
Sequencing vector: M13, 93%
Sequencing vector: plasmid, 7%
Chemistry: Dye-terminator Big Dye, 7% of reads
Chemistry: Dye-terminator Big Dye, 7% of reads
Assembly program: Phrap, version 0.990319
Consensus quality: 179082 bases at least Q40
Consensus quality: 183120 bases at least Q30
Consensus quality: 185448 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 190318; sum-of-contigs
Quality coverage: 3.99 in Q20 bases; sum-of-contigs
Quality coverage: 4.34 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1279: contig of 1279 bp in length
* 1280 1379: gap of unknown length
* 1380 3544: contig of 2165 bp in length
* 3545 3644: gap of unknown length
* 3645 5441: contig of 1797 bp in length
* 5442 5541: gap of unknown length
* 5542 8421: contig of 2880 bp in length
* 8422 8521: gap of unknown length
* 8522 10096: contig of 1575 bp in length
* 10097 10196: gap of unknown length
* 10197 11938: contig of 1742 bp in length
* 11939 12038: gap of unknown length
* 12039 14708: contig of 2670 bp in length
* 14709 14808: gap of unknown length
* 14809 17876: contig of 3068 bp in length
* 17877 17976: gap of unknown length
* 17977 21534: contig of 3558 bp in length
* 21535 21634: gap of unknown length
* 21635 24615: contig of 2981 bp in length
* 24616 24715: gap of unknown length
* 24716 28301: contig of 3586 bp in length
* 28302 28401: gap of unknown length
* 28401 35969: contig of 7568 bp in length
* 35970 36059: gap of unknown length
* 36070 42473: contig of 6404 bp in length
* 42474 42573: gap of unknown length
* 42574 49752: contig of 7179 bp in length
* 49753 49852: gap of unknown length
* 49853 57604: contig of 7752 bp in length
* 57605 57704: gap of unknown length
* 57705 67088: contig of 9384 bp in length
* 67089 67188: gap of unknown length
* 67189 77656: contig of 10468 bp in length
* 77657 77756: gap of unknown length
* 77757 92143: contig of 14387 bp in length

```

FEATURES

source

ORIGIN

[illegible]

OY	2921	CACCACTTTCATTTTCAGCTCCCTGTAGGACATCTCTGAATTTAGACACAGAAAATGCC	2980
Db	41276	AACCACTTTCATTTTCAGCTCCCTGTAGGACATCTCTGAATTTAGACACAGAAAATGCC	41217
OY	2981	TCGTAGTCACTGAATCCAGAAAAGCTCTTACCTTTAGACACAAGGAGCTTCACCA	3040
Db	41216	TCGTAGACACTGAATCCAGAAAAGCTCTTCACCTTTAGACACAAGGAGCTTCACCA	41157
OY	3041	CTGGAC-AAAAGAGAAAGATAGGGGTAGTACCAAGAACTCTCTTCCACAGTCAGT	3099
Db	41156	CAGGACAAAAGAGAAAGATAGGGGTAGTACCAAGAACTCTCTTCCACAGTCAGT	41057
OY	3100	TATGATTTTTGCTGTAGATCATCTCTTATGCTTCCACTTGGGTACATAGCAGGGGG	3159
Db	41096	TATGTTTTTTCCTATCGATCATCTTTC- TGTCCTTCCACCTCGTGCTACATGACGGGG	41038
OY	3160	TCACGAGCTTGTTCAGAAAAAGACAGAGACATGAAGCTTCTTTCAGAACTGAGTGC	3219
Db	41037	TCATGAGCTTGTTCAGAAAAAGACAGAGACACGAAGCTTCTGTTTGGAAACTGAGTGC	40978
OY	3220	TGTCAACCCCAACTGTGTGAGCTTAAATGCTGTCCCCCTTTCTAATTTATCTCCCCATA	3279
Db	40977	TGCACACCCCAACTGTGTGAGCTTAAATGCTGTCCCCCTTTCTAATTTATCTCCCCATA	40919
OY	3280	TCACCTCTCTCATTCACATCATTCGAATCTGCGCTCATGAGAGACTGCGCTCTTACAT	3339
Db	40918	TCATCTCTCTCATTCGAATCATTCGAATCTGCTCTCTGAGAGAGATGCTGCTCTTACAT	40859
OY	3340	TCATTTAAGAGCAAGGGGACATGACAGGCAATTTCTTCCAGATTGAACCTGCTATAGAC	3399
Db	40858	TCATTTAAGAGGCGACAGAAAGATGAAGCATTTCTTCCAGATTGAACCTGCTATAGGCG	40799
OY	3400	CAGTTCTTGTGTTCACCTACTTTCAAATTTATCTTCTTGGCTATCTGGAAGAGTCT	3459
Db	40798	CGTTTATTTGTTCACCTACTTTCAAATTTATCTTCTTGGCTATCTGGAAGAGACC	40799
OY	3460	AAGAGATATTAAGATGCGCCAA-----TAATTAAGAGTGTTCATGAGAAAGTAT	3512
Db	40738	AAGGAATTAAGATGATGCTCCAACTGAATTAATTAAGAGTGTTCATGAGAAAGTAT	40679
OY	3513	TACAAAGATGSCAGAGTTAA-----GGGTGAGANTCCTAACACAGA	3554
Db	40678	TACAAAGATGSCAGAGGTTAAAGGAAAGATGAGGCAATGGAAGAACCTCAACACAGA	40619
OY	3555	ATACATAGGGGAGCACTACTCTCTCCCTAGCTGAAACGACACAGGAAAGAGCAAGTTAC	3614
Db	40618	ATACATAGGGGAGCACTACTCTCTCTCCCTAGCTGAAAGGACACAGGAAAGAGCAAGTTAC	40559
OY	3615	CATTGTGCGCATAGCCATAGCTGTAGCCATTAAGGTGGGAAGCATGAGCAGGCAAGTGG	3674
Db	40558	CATTGTGCGCATAGCCATAGCTGTAGCCATTAAGGTGGGAAGGACACAGCAGGCAAGTGG	40499
OY	3675	AGAAAGCTGCGTGGCCAAAGCAGCAGCACAAGGCTGATTAAGTTAGTATGATGTGCTCC	3734
Db	40498	AGAAAGCTGCGCAGCCAAACACACACACACAGGCTATATGTTTGGATCTGTGTCCC	40439
OY	3735	CACCAAAATCTAGTGTGATTTGAATTTCCATGTTTGGAGAAAGGCGCTTGTGGAGATG	3794
Db	40438	TACCAAAATCTAGTGTGATTTGAATTTCCATGTTTGGAGGAAGGCGCTGTGGG--GATG	40381
OY	3795	ATTATTAAGATCAGGGGATGTTTTGCAATGAATGTTTTAAACCAATCCCTTTGGTAT	3854
Db	40380	ATTATTAAGATCAGGGGATGTTTTGCAATGAATGTTTTAAACCAATCCCTTTGGTAT	40321
OY	3855	GTTGTGTGATCTGACGAGTCTCATGAATCTAGTGTTTAAAGCGTGTAGACCTC	3914
Db	40320	AATGTTTGTATCTGATTAAGTTCATGAATCTAATTTGTTTAAAGGTGTAGACCTC	40261
OY	3915	CTCC--CTCTCTTATCTCTGCTCTCACCATGTGAGACGCTGCTGCCCTTTGCTCT	3972
Db	40260	CTCCCTCTCTCTGTGCTCTGCTCTCACCATGTGAGAGTGCCTTGCCTGCCCTTTGCTCT	40201

QY	3973	CCACAGGATTGGAAAGCTTCTGAGGCTCCCGCAAGACGAAAGCTGATAGCTCTGT	4032
Db	40200	TCACGAGATTGGAAAGCTTCTGAAGCTCCCTAAAGCAGAACTGATGCTTCTGT	4014
QY	4033	AACGTCTGTAAAGCTATTAGCCAGTAAACCCATTTCTCTATTAATTTCCAGTCTAG	4092
Db	40140	ATAGCTGTAGGCGATAGGCCAATTAACCTGTCTTTCATATTAATTTCCAGTCTAG	4008
QY	4093	GTATTTCTTTTAGCAATTGAGAAATGAATCTAATATACACAGACAGAGCCAGAGATNGA	4152
Db	40080	GTATTTCTTTTATTCAAAGCAAGAAAGAACTAAATACGACGACAGAGCCAGAGATNGA	4002
QY	4153	AATCCCAAGTGTCTTCTGCTGTCTTCCAGTCTCTGCTGTGTCTTCCAGTGTCTCAA	4212
Db	40020	AATCCCAAGTCACTCTCTGTGTCTTCCAGTCTCTGCTGTGTCTTCCAGTGTCTCAA	3996
QY	4213	TTCCACCAAAACGAAATTAAGAAATCCCATGTGTGTATCATGAAGCCACTCTC	4272
Db	39960	TTTCCTTTGAAACCAAGAAATTAAGAAATCTCAATGTGTGTACAGGAAGCCACTCAC	3990
QY	4273	TTGGGATGTCAACAGGATTAAGAAATGGAAGCAAAATCTCATATGTAAATGAGCTA	4332
Db	39900	TTGGGATGTCAACAGGATTAAGAAATGGAAGCAAAATCTCATATGTAAATGAGCTA	3984
QY	4333	TCCCTTCACCTCTGTGTATCTCTCATATTCCTGCGGCTTCTCTCT-----ATCTGATG	4386
Db	39840	TTCCCTTACCTCTGTATATCTCTCATATTCCTGAGGCTGTCTGCTGCAAGGTGATNG	3978
QY	4387	ATCCCTGTCTCATTTACGCTCTATACACATCTTTAATGTTGGCTGTCTTCTCTACT	4446
Db	39780	ATCTCTGTCTCATTTACGCTCTATGAGCTATCTTAAATGTTGGCAGTCTTCTCTACT	3972
QY	4447	GTCACTTTTATGACGAATGTTTCATTTGTTTAAATATGATATGATAATTAATGTATTT	4506
Db	39720	GTCACTTTTATGACGAATGTTTCATTTGTTTAAATATGATATGATAATTAATGTATTT	3966
QY	4507	TTAAAGAAACATATGTAATTTTGTTTGAATATATAGTTTGGCTGTATTAATAAAGCATG	4566
Db	39660	TAAAGAAAGAAACATATTTATTTGTTGAATATATAGTTTGGCTGTATTAAGAAAGCATG	3960
QY	4567	AAGAAGAAATATCTTAAACAGAAAGATATAGTTTGGCTGTGGGTCACATGAGTCTGAAT	4626
Db	39600	AAGTATGAATATCTTAAACAGAAAGATATAGTTTGGCTGTGGGTCACATGAGTCTGAAT	3954
QY	4627	CTACAGATTCAACAACTACGAGAGAACTTTTCCAAAATAAAGGTGTGGCGAGTTG	4686
Db	39540	CTGCAAAATTCACAAACCAACAGTGGAACTATTTCCAAAATAAAGGTATGTGTGAGATTG	3948
QY	4687	TGTATGTATCTGAACAGGTACAACTGTGATTTCTTGTCAATATTTCTGAAAACTTCAAA	4746
Db	39480	TGTATGTATCTGAACAGGTACAACTGTGATTTCTTGTCAATATTTCTGAAAACTTCAAA	3942
QY	4747	TATAACAGAACTTATATAGCAATTTGCAATTTGTCAATATTTCTTAAATTAATTTAAATGA	4806
Db	39420	TATAACAGAAATTTATATATCAATTTTGCAATTTCAATGATTTCTTAAATTAATTAATGA	3936
QY	4807	TTTAATGTATCTGGAGAAAGTGCATGAGTATATACAAATACATATATATAGAAATTTG	4866
Db	39360	TTTAACATATCTGGAGAAAGTGTGATAG-GTATATTAACAAATACCAATATATAGAAATAG	3930
QY	4867	AGCATCTGCAATTTTGTGTGTGTCTGTGGGGTTCTGAAAAAGAAATCCCTGTAAATACAA	4926
Db	39300	AGCATCTGCAATTTTGTGTGTGTCTGTAAAGTTCTTGAAAAAATCCCTGTGAATACCCAG	3924
QY	4927	AAATGACACTTTCGAGATCTGAACTGAGAGCTCCAAAGCATCATACAGAAATTTCCAA	4986
Db	39240	AAATGACATCTTTCGAAATCTGAAATGTAGAGCTCCAAAGCGTGCACATACAGAAATTTCCAA	3918
QY	4987	AAATTTGCTGTCCCAAGTTCTTAAGAGTTGGCTCATCTCTGTGATCTCAATAGTTTCC	5046
Db	39180	AAATTTGGGGCT-CCAGTCTCTTAAGAGTTGGCTCTCATCTTGTGATCTTGCAATGATTTTC	3912
QY	5047	CAGGACATTTGCAATTCAGTCTTATGAAAAAGACGAGGGGAAAGAGAGCTTGTCTC	5106

QY 7255 TTGATCTTAATGTTATCCCAATGACAGAGAACTTGTGGAGTAGAGAGTCA 7314
 Db 36929 TTTATCTTAATGTTATCCCAATGACAGAGAACTTGTGGAGTAGAGAGTCA 36870
 QY 7315 GGCCTCAGAGTCAACAGAACTGATTTTCAACTGATTTGAGACCCCACTTTGAT 7374
 Db 36869 GGCCTCAGAGTCAACAGAACTGATTTTCAACTGATTTGAGACCCCACTTTGAT 36811
 QY 7375 AGGTGACTTATTTCTGTGAGTCTCTG-ATCTCTCTCTTTTAATGAGACAGTAATATCC 7433
 Db 36810 AAGTGACTTATTTCTGTGAGTCTCTG-ATCTCTCTCTTTTAATGAGACAGTAATATCC 36751
 QY 7434 CACATGAGAGGTGTGGGAGAGATCAGAGATCAACAGCTGTGATCACTCTGGTTTC 7493
 Db 36750 CATACGAGAGGTGTGGGAGAGATCAGAGATCAACAGCTGTGATCACTCTGGTTTC 36691
 QY 7494 TGTTCACAGGTGATCAGACTGGGGTTTCTGAGCATGATTCACCAATCCAGTCTTGG 7553
 Db 36690 TGTTCACAGGTGATCAGACTGGGGTTTCTGAGCATGATTCACCAATCCAGTCTTGG 36631
 QY 7554 TACGAACTGACACCAATCAACGACGTGAGAGACTCTTGTACAGACAGACCTTGAG 7613
 Db 36630 TACGAACTGACACCAATCAACGACGTGAGAGACTCTTGTACAGACAGACCTTGAG 36571
 QY 7614 CTTACGAGGGGCTGAGTGTGATGTTCCCTGTGCGCTGACAGAGAAAGCGGTTGTGT 7673
 Db 36570 CTTACGAGGGGCTGAGTGTGATGTTCCCTGTGCGCTGACAGAGAAAGCGGTTGTGT 36511
 QY 7674 CTGGCTCTCTGAGGCTGCGCATGCGAGAGACGCTGTCTCATCTACATCCTCAACTGTG 7733
 Db 36510 CTGGCTCTCTGAGGCTGCGCATGCGAGAGACGCTGTCTCATCTACATCCTCAACTGTG 36451
 QY 7734 CGCGGCGCACTTCTCTTCTTGAAGGCGCACTTATATGTTGCGCTTACGCTCATCA 7793
 Db 36450 CGCGGCGCACTTCTCTTCTTGAAGGCGCACTTATATGTTGCGCTTACGCTCATCA 36391
 QY 7794 TATCCGACCATCCATCTCCAAATCCTCAGTCTGTGATGACCTTCCCTACTTTTAG 7853
 Db 36390 TATCCGACCATCCATCTCCAAATCCTCAGTCTGTGATGACCTTCCCTACTTTTAG 36331
 QY 7854 CCTAAGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGTCATCTGTGGCCATCTG 7913
 Db 36330 CCTAAGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGTCATCTGTGGCCATCTG 36271
 QY 7914 GTACACATGCGCGCGCCCAATACCTGTCATCTGTGATGACCTTCCCTACTTTTAG 7973
 Db 36270 GTACACATGCGCGCGCCCAATACCTGTCATCTGTGATGACCTTCCCTACTTTTAG 36213
 QY 7974 GTCCCTGTGCGAGATCTCTGAGTGAATGTTCTGTGACTTCTGTAGTGTCTGA 8033
 Db 36212 GTCCCTGTGCGAGATCTCTGAGTGAATGTTCTGTGACTTCTGTAGTGTCTGA 36153
 QY 8034 TTTCTGTTGTTGTTGAACGTCAGATTTCAATACATCGCGTGTGTTTTTTTATGTGT 8093
 Db 36152 TTTCTGTTGTTGTTGAACGTCAGATTTCAATACATCGCGTGTGTTTTTTTATGTGT 36096
 QY 8094 GGTCTCTGTGTGTGTCAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8153
 Db 36095 GGTCTCTGTGTGTGTCAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 36037
 QY 8154 GCCCGTGAACAGGCTGTGATGACCATCTCTCTCAAGTGTGTGTCTCTCTGTGTGTGT 8213
 Db 36036 NNN 35977
 QY 8214 CCTGCGCTT 8222
 Db 35976 NNNNNNNNTT 35968

RESULT 7
 AC020568
 LOCUS AC020568 192618 bp DNA linear HTG 17-AUG-2000

DEFINITION Homo sapiens chromosome 20 clone RP11-206C1, WORKING DRAFT
 SEQUENCE, 24 unordered pieces.
 AC020568
 AC020568.4 GI:9280789
 HTG, HTGS, PHASE1, HTGS, DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 192618)
 Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 192618)
 Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (04-JAN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Jul 19, 2000 this sequence version replaced gi:9211362.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H.NH0206C01
 ----- Summary Statistics -----
 Sequencing vector: M13, 93%
 Sequencing vector: plasmid, 7%
 Chemistry: Dye-terminator Big Dye, 7% of reads
 Chemistry: Dye-terminator Big Dye, 7% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 179082 bases at least Q40
 Consensus quality: 183120 bases at least Q30
 Consensus quality: 185448 bases at least Q20
 Insert size: 20500; agarose-fp
 Insert size: 190318; sum-of-contigs
 Quality coverage: 3.99 in Q20 bases; agarose-fp
 Quality coverage: 4.34 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 24 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1279: contig of 1279 bp in length
 * 1280 1379: gap of unknown length
 * 1380 3544: contig of 2165 bp in length
 * 3545 3644: gap of unknown length
 * 3645 5441: contig of 1797 bp in length
 * 5442 5541: gap of unknown length
 * 5542 8421: contig of 2880 bp in length
 * 8422 8521: gap of unknown length
 * 8522 10096: contig of 1575 bp in length
 * 10097 10196: gap of unknown length
 * 10197 11938: contig of 1742 bp in length
 * 11939 12038: gap of unknown length
 * 12039 14708: contig of 2670 bp in length
 * 14709 14808: gap of unknown length
 * 14809 17876: contig of 3068 bp in length
 * 17877 17976: gap of unknown length
 * 17977 21534: contig of 3558 bp in length
 * 21535 21634: gap of unknown length
 * 21635 24615: contig of 2981 bp in length
 * 24616 24715: gap of unknown length
 * 24716 28301: contig of 3586 bp in length
 * 28302 28401: gap of unknown length
 * 28402 35969: contig of 7568 bp in length
 * 35970 36069: gap of unknown length

*	3660	42477	contig of 6404 bp in length
*	42474	42577	gap of unknown length
*	42574	42952	contig of 7179 bp in length
*	49753	49852	gap of unknown length
*	49853	57604	contig of 7752 bp in length
*	57604	57705	gap of unknown length
*	57705	57088	contig of 9384 bp in length
*	67089	67189	gap of unknown length
*	67189	77655	contig of 10466 bp in length
*	77657	77755	gap of unknown length
*	77757	92143	contig of 14387 bp in length
*	92144	92243	gap of unknown length
*	92243	109240	contig of 16997 bp in length
*	109241	109340	gap of unknown length
*	109341	124285	contig of 14945 bp in length
*	124285	124385	gap of unknown length
*	124386	138608	contig of 14223 bp in length
*	138609	138708	gap of unknown length
*	138709	155882	contig of 15174 bp in length
*	155883	153982	gap of unknown length
*	153983	165994	contig of 16012 bp in length
*	165995	170094	gap of unknown length
*	170095	192618	contig of 22524 bp in length

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="caxon:9606"
/chromosome="20"
/clone="RP11-206C1"
1. 1279
misc_feature
/note="assembly_name: Contig19
1380. 3544
/note="assembly_name: Contig21
3645. 5441
/note="assembly_name: Contig22
5542. 8421
/note="assembly_name: Contig23
8522. 10096
/note="assembly_name: Contig24
10137. 11938
/note="assembly_name: Contig25
12039. 14708
/note="assembly_name: Contig26
14809. 17876
/note="assembly_name: Contig27
17977. 21534
/note="assembly_name: Contig28
21635. 24615
/note="assembly_name: Contig29
2476. 28301
/note="assembly_name: Contig30
28402. 33569
/note="assembly_name: Contig31
clone end: SP6
vector_side:left"
36070. 42473
/note="assembly_name: Contig32
42574. 49752
/note="assembly_name: Contig33
49853. 57604
/note="assembly_name: Contig34
57705. 67088
/note="assembly_name: Contig35
67189. 77656
/note="assembly_name: Contig36
77757. 92143
/note="assembly_name: Contig37
92244. 109240
/note="assembly_name: Contig38
109341. 124285
/note="assembly_name: Contig39
124386. 138608

```

```
misc_feature      /note="assembly_name:Contig40"  
                  138709..153882  
note="assembly_name:Contig41"  
misc_feature      153983..169394  
note="assembly_name:Contig42"  
misc_feature      170095..192618  
note="assembly_name:Contig43"
```

Query Match	52.1%	Score 4492.2;	DB 2;	Length 192618;
Best Local Similarity	99.6%;	Pred. No. 0;		
Matches 4555; Conservative	0;	Mismatches 13;	Indels 5;	Gaps 5;

ORIGIN

QY	1	TGTAATGAAGCCATGTCACTTTAATACCAAAAC	CAGAAAAGATATATCAAAAAAGAAAAC	60
Db	53037	TGTAATGAAGCCAAATGTCACTTTAATACCAAAAC	CAGAAAAGATATATCAAAAAAGAAAAC	53096
QY	61	TATAGACACAGTACCACTGATGAATATATCATGT	CAGAAATCCCAACAAATATCTAGCTAAC	120
Db	53097	TATAGACACAGTACCACTGATGAATATATCATGT	CAGAAATCCCAACAAATATCTAGCTAAC	53156
QY	121	CCAATCCAAACGACATATCAAGAAATATCCACAT	TGTCAAGTGGGTTCTATACCAAGG	180
Db	53157	TCAATCCAAACGACATATCAAGAAATATATCCACAT	TGTCAAGTGGGTTCTATACCAAGG	53216
QY	181	GTGCAGATATGTTTACATATCAACAAGTCAATTA	ATGTGATATCATCATTAACAGATTA	240
Db	53217	GTGCAGATATGTTTACATATCAACAAGTCAATTAAT	-TGATATCATCATTAACAGATTA	53275
QY	241	AAAACAAAATTCACATGATATCTCAATTAATGA	TGCTGAAAAAGATTTGCAAAATCTAAC	300
Db	53276	AAAACAAAATTCACATGATATCTCAATTAATGATAT	CTGAAAAAGATTTGCAAAATCTAAC	53335
QY	301	ATTTCTTATGATTAATAAACCCTTCAGCAAAAT	CCACATAGAAAAGACATACCTTAATGTAA	360
Db	53336	ATTTCTTATGATTAATAAACCCTTCAGCAAAAT	CCACATAGAAAAGACATACCTTAATGTAA	53395
QY	361	TAAAGCCATATATATGACGAGCCACAGCAAACT	TATCTGAATGGGAAAAAGTTGAAAA	420
Db	53396	TAAAGCCATATATATGACGAGCCACAGCAAACT	TATCTGAATGGGAAAAAGTTGAAAA	53455
QY	421	CATTGTCCCTGAGAACTTGAAACAAGATGTAAT	TACCACTTCATATCAACAT	480
Db	53456	CATTGTCCCTGAGAACTTGAAACAAGATGTAAT	TACCACTTCATATCAACAT	53515
QY	481	AGTAGTGGAAAGTTTATGACGAGCAATCAGACA	GAGAAAGAAATCAAGGCAACCCAAT	540
Db	53516	AGTAGTGGAAAGTTTATGACGAGCAATCAGACA	GAGAAAGAAATCAAGGCAACCCAAT	53575
QY	541	CAATTAAGAGAAATCTMAACTGTCCCTGTTCACT	GATGATATGATTTGTTATCTAGAAAA	600
Db	53576	CAATTAAGAGAAATCTMAACTGTCCCTGTTCACT	GATGATATGATTTGTTATCTAGAAAA	53635
QY	601	CCCTTAAAGACTCATCCGAAAAGCTCCTGAACT	GTATACATAAATTCAGTAAAGTTTCAGG	660
Db	53636	CCCTTAAAGACTCATCCGAAAAGCTCCTGAACT	GTATACATAAATTCAGTAAAGTTTCAGG	53695
QY	661	ATACAAACTAATSTACAAATACAGAGACGTATAT	ACCAACAGTGACCAAGCTGA	720
Db	53696	ATACAAACTAATSTACAAATACAGAGACGTATAT	ACCAACAGTGACCAAGCTGA	53755
QY	721	GAATCAATCAATCAAGAACTCAACAACCTTTTAC	ATATAGCTGTATAAAAAATACTTAAGAATATT	780
Db	53756	GAATCAATCAATCAAGAACTCAACAACCTTTTAC	ATATAGCTGTATAAAAAATACTTAAGAATATT	53815
QY	781	CTTATCCCAAGAGGTGAAGGACCTCTTACAAAG	AAAACTTACAAAAACACAGCTGCATCATTA	840
Db	53816	CTTATCCCAAGAGGTGAAGGACCTCTTACAAAG	AAAACTTACAAAAACACAGCTGCATCATTA	53875
QY	841	GATATACCAAAACAAGTGGAAACAATCCCATGCT	CACTGATGGGTGAATTCATATATTGTG	900
Db	53876	GATATACCAAAACAAGTGGAAACAATCCCATGCT	CACTGATGGGTGAATTCATATATTGTG	53935

[illegible]

Db	55014	CATTGATCCAGCAGCATCCATTAAATGTATATAATATATATATATATACATGAAT	55073
QY	2041	ACAACCTAGCCATTAATAAATAAATAATGATGACATTCACAGCAATCTAGATGAAATGG	2100
Db	55074	ACAACCTAGCCATTAATAAATAAATAATGATGACATTCACAGCAATCTAGATGAAATGG	55133
QY	2101	AGACCTTTATTTCTAAGTGGGGTAACTCAGAGATGAAAAACAAATCATATGTTCTCAC	2160
Db	55134	AGACCTTTATTTCTAAGTGGGGTAACTCAGAGATGAAAAACAAATCATATGTTCTCAC	55193
QY	2161	TTACAAGTGGGGGCTAAGCTGTAGAGACACGAAGGATGGAATGATATATGAATCTGG	2220
Db	55194	TTACAAGTGGGGGCTAAGCTGTAGAGACACGAAGGATGGAATGATATATGAATCTGG	55253
QY	2221	GGACTTGAAGGGAAGATGGAAGAGGCGAGGATTAAGAATCACAAATGGGTACAGT	2280
Db	55254	GGACTTGAAGGGAAGATGGAAGAGGCGAGGATTAAGAATCACAAATGGGTACAGT	55313
QY	2281	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTCACCTAAGAATCTATC	2340
Db	55314	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTCACCTAAGAATCTATC	55373
QY	2341	CATGAAGCAACAACCACTGTTCCCAAAATCCCAATGAAATTAATAATATATTA	2400
Db	55374	CATGAAGCAACAACCACTGTTCCCAAAATCCCAATGAAATTAATAATATATTA	55433
QY	2401	ATGATTTTAAATTCACAGAAATTTAAAAAAGTTCACCTGTTCAGAGTTTATATATGAAGTA	2460
Db	55434	ATGATTTTAAATTCACAGAAATTTAAAAAAGTTCACCTGTTCAGAGTTTATATATGAAGTA	55493
QY	2461	AGAAATGAAGAGTGTAGCAAGTGTGAGCTCTGAGCAATGGGACTTGAATTTCACTTG	2520
Db	55494	AGAAATGAAGAGTGTAGCAAGTGTGAGCTCTGAGCAATGGGACTTGAATTTCACTTG	55553
QY	2521	CATACACTTCTCTGGCATTTTGAAAAAGTATACACATGAATATATTCACCTATGATA	2580
Db	55554	CATACACTTCTCTGGCATTTTGAAAAAGTATACACATGAATATATTCACCTATGATA	55613
QY	2581	AAGAAAAATCAAAAAAATGTGTCAAGGCACTTGTCAAGCTTGAAATGGTCCCATGATCTAC	2640
Db	55614	AAGAAAAATCAAAAAAATGTGTCAAGGCACTTGTCAAGCTTGAAATGGTCCCATGATCTAC	55673
QY	2641	TTTTTCATTTTGATATATAAGCCTCATATATGATGTTCACTTGTCTTAATGTGATGCTTGA	2700
Db	55674	TTTTTCATTTTGATATATAAGCCTCATATATGATGTTCACTTGTCTTAATGTGATGCTTGA	55733
QY	2701	GCCCATATTTGATTTTAAATCAGGACAGCAATTACTTAACAGGAAGTTGAACAAGATGG	2760
Db	55734	GCCCATATTTGATTTTAAATCAGGACAGCAATTACTTAACAGGAAGTTGAACAAGATGG	55793
QY	2761	GACGTGATAGGAGGCTTAAATGTATATGGAATATGGGACAGAGGCCAAGATCATCTCAG	2820
Db	55794	GACGTGATAGGAGGCTTAAATGTATATGGAATATGGGACAGAGGCCAAGATCATCTCAG	55853
QY	2821	TTAGGATTTGTGTCTCAAAATACCTCTGGCCTCTGATTTTGCCATAGTCTCTACAGAA	2880
Db	55854	TTAGGATTTGTGTCTCAAAATACCTCTGGCCTCTGATTTTGCCATAGTCTCTACAGAA	55913
QY	2881	ATTAACAGACTGTCTCAGACTCTTGTGTAAGCTTGATTTGCTCACAGCTTTCAATTCAGT	2940
Db	55914	ATTAACAGACTGTCTCAGACTCTTGTGTAAGCTTGATTTGCTCACAGCTTTCAATTCAGT	55973
QY	2941	CCTGAAGGCATCTCCTGAATTAAGCAACAGAAAAAGTCCCTGGAAGTCACTGAATCCCA	3000
Db	55974	CCTGAAGGCATCTCCTGAATTAAGCAACAGAAAAAGTCCCTGGAAGTCACTGAATCCCA	56033
QY	3001	GAAAGGCTCTTACCTTTAGCAACAAGGAGGTCTTCAACCACTGGAACAAAGAAAGCAT	3060
Db	56034	TAAAGGCTCTTACCTTTAGCAACAAGGAGGTCTTCAACCACTGGAACAAAGAAAG-CAGAT	56092
QY	3061	AAGGTAATGACAAAGACTCTCTTTCTTCCACAGTCACTTAATTTTGTCTGTAAATC	3120

Db 56093 AAGGTAAGTA-CAGACACTCTCTTCCACAGTCAAGTATTTTGTGTAAGATC 56151
 Qy 3121 ATGTCTTATGCTTCCACCTTGTGTCTACATGACGGGGGTCCAGAGCTTTTTCAGAAA 3180
 Db 56152 ATGTCTTATGCTTCCACCTTGTGTGTACATGACGGGGGTCCAGAGCTTTTTCAGAAA 56211
 Qy 3181 AGACAGAGACATGAAAGCTTCTTTCAGAAATGAGTGTCTGACACCCAACTGTGTAG 3240
 Db 56212 AGACAGAGACATGAAAGCTTCTTTCAGAAATGAGTGTCTGACACCCAACTGTGTAG 56271
 Qy 3241 CTCTAAATGAGTGTCCCTCTTAATTTATCTCCCATATCACTCTCTTCAATTCATCA 3300
 Db 56272 CTCTAAATGAGTGTCCCTCTCTTAATTTATCTCCCATATCACTCTCTTCAATTCATCA 56331
 Qy 3301 TTCAATCTGCTTCATGAGAGACGTGCTCTTATCATTTCAATTAACAGACAGGGAC 3360
 Db 56332 TTCAATCTGCTTCATGAGAGACGTGCTCTTATCATTTCAATTAACAGACAGGGAC 56391
 Qy 3361 ATGACAGCATTTCTTCCAGAGTTGAACGTCTATAGCCAGTTCTTTTTCATTTAC 3420
 Db 56392 ATGACAGCATTTCTTCCAGAGTTGAACGTCTATAGCCAGTTCTTTTTCATTTAC 56451
 Qy 3421 TTTTCAATTTATCTCTTGTGCTATCTGGAAGGTCTAAGAAATATATGATGAGCCCA 3480
 Db 56452 TTTTCAATTTATCTCTTGTGCTATCTGGAAGGTCTAAGAAATATATGATGAGCCCA 56511
 Qy 3481 ATATTAAGAGTGTCTCATGAGAAAGTATTTTCAAAAGATGACAGAGTTAAGGCTCAG 3540
 Db 56512 ATATTAAGAGTGTCTCATGAGAAAGTATTTTCAAAAGATGACAGAGTTAAGGCTCAG 56571
 Qy 3541 GATCTTAAGACATATCATGAGGAGCATTAATCTTCTCCCTTGTGCTGAACGAGACAG 3600
 Db 56572 GATCTTAAGACATATCATGAGGAGCATTAATCTTCTCCCTTGTGCTGAACGAGACAG 56631
 Qy 3601 GAAGAGACATTAATCATTTGTGCTGACATAGCCATAGCTTATGAGGATGAGAGCAT 3660
 Db 56632 GAAGAGACATTAATCATTTGTGCTGACATAGCCATAGCTTATGAGGATGAGAGCAT 56691
 Qy 3661 GAGAGAGCATGAGAGAGCCCTGCTGAGCCACGACAGCAGCAGAGCTGATATAGTT 3720
 Db 56692 GAGAGAGCATGAGAGAGCCCTGCTGAGCCACGACAGCAGCAGAGCTGATATAGTT 56751
 Qy 3721 TGGATCTGTGTTCCCAACCAAAATCTCATTTGATTTTCCAAATGTTGAGAGAGG 3780
 Db 56752 TGGATCTGTGTTCCCAACCAAAATCTCATTTGATTTTCCAAATGTTGAGAGAGG 56811
 Qy 3781 CCTTGTGAGAGATATTAATGATCAGGGGATGTTTGTGATGAATGTTTAAACACAT 3840
 Db 56812 CCTTGTGAGAGATATTAATGATCAGGGGATGTTTGTGATGAATGTTTAAACACAT 56871
 Qy 3841 CCCCCTTTGATTTGTTGTGTGATGATGACAGAGTTCTCATGAATCTAGTTGTTTAAA 3900
 Db 56872 CCCCCTTTGATTTGTTGTGTGATGATGACAGAGTTCTCATGAATCTAGTTGTTTAAA 56931
 Qy 3901 GCGGTATGACACTCTCTCTCTCTTAATCTCTGCTCTCACCATGTGAGAGCCCTCGCTC 3960
 Db 56932 GCGGTATGACACTCTCTCTCTCTTAATCTCTGCTCTCACCATGTGAGAGCCCTCGCTC 56991
 Qy 3961 CCCCCTTTGCTTTCACAGAGATGGAAGCTTCTGAGGCTCCCAAGAGAGAGAGCTGC 4020
 Db 56992 CCCCCTTTGCTTTCACAGAGATGGAAGCTTCTGAGGCTCCCAAGAGAGAGAGCTGC 57051
 Qy 4021 TATGCTTCTTGTACAGTCTGTAGAGCTATTAAGCAGTTAAACCAATTTCTTCAATAAT 4080
 Db 57052 TATGCTTCTTGTACAGTCTGTAGAGCTATTAAGCAGTTAAACCAATTTCTTCAATAAT 57111
 Qy 4081 TCCCAAGTCTCAGAGATTTCTTTTAAAGCAATTTGAGAAATTAATACAGACAGAGAG 4140
 Db 57112 TCCCAAGTCTCAGAGATTTCTTTTAAAGCAATTTGAGAAATTAATACAGACAGAGAG 57171
 Qy 4141 CCAGAGATGGAATTCAGAGGCTTCTCTGTCTTCCAGTCTCTCTGTGCTGCTC 4200
 Db 57172 CCAGAGATGGAATTCAGAGGCTTCTCTGTCTTCCAGTCTCTCTGTGCTGCTC 57231

Qy 4201 CCAGTGTCTCAATTTCCACCAAGAAATTAAGAAATCCCATCTGATGTGTACATA 4260
 Db 57232 CCAGTGTCTCAATTTCCACCAAGAAATTAAGAAATCCCATCTGATGTGTACATA 57291
 Qy 4261 GAAGCCACTCTTGGGATGTCAAAACGATTAAGAAATGGAAGCAATCTCATGG 4320
 Db 57292 GAAGCCACTCTTGGGATGTCAAAACGATTAAGAAATGGAAGCAATCTCATGG 57351
 Qy 4321 TAAATGAGACTATCCCTCTGACCTTCTTGTATCTCTTAATTCCTGGGGCTTCTATC 4380
 Db 57352 TAAATGAGACTATCCCTCTGACCTTCTTGTATCTCTTAATTCCTGGGGCTTCTATC 57411
 Qy 4381 TGAATGATCCCTGTCTCATTTGACCTCTATCACTATCTTAATGTTGGCTGTCTTC 4440
 Db 57412 TGAATGATCCCTGTCTCATTTGACCTCTATCACTATCTTAATGTTGGCTGTCTTC 57471
 Qy 4441 TCTACTGTCACTTTTAAAGCAAAATGTTGCAATTTGTTAAAGATGCAATGAATTAAT 4500
 Db 57472 TCTACTGTCACTTTTAAAGCAAAATGTTGCAATTTGTTAAAGATGCAATGAATTAAT 57531
 Qy 4501 GTAAATTTAAAGCAATATGATTTTGTGTAAGATATATGTTGCTGATCTAATPAA 4560
 Db 57532 GTAAATTTAAAGCAATATGATTTTGTGTAAGATATATGTTGCTGATCTAATPAA 57591
 Qy 4561 GACATGAAGAGA 4573
 Db 57592 GACATGAAGAGA 57604

RESULT 8
 AC139482/c 165490 bp DNA linear HTG 04-FEB-2003
 LOCUS Homo sapiens chromosome 5 clone RP11-583F24, WORKING DRAFT
 DEFINITION
 SEQUENCE: 12 unordered pieces.
 AC139482
 AC139482.1 GI:28201512
 HTG; HTG_PHASE1; HTG_DRAFT; HTG_ACTIVEPIN.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

Project Information
 Center Project Name: 1518520
 Center clone name: RPC1-11_583F24

 Summary Statistics
 Consensus quality: 147035 bases at least Q40
 Consensus quality: 155210 bases at least Q30
 Consensus quality: 158567 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 164390; sum-of-contigs estimation
 Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
 Quality coverage: 4.19 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1249: contig of 1249 bp in length
 * 1350 1349: gap of unknown length
 * 1350 1350: contig of 2331 bp in length
 * 3780 3780: gap of unknown length
 * 3781 3668: contig of 2888 bp in length
 * 6669 6768: gap of unknown length
 * 6769 10866: contig of 4098 bp in length
 * 10867 10966: gap of unknown length
 * 10967 17019: contig of 6053 bp in length
 * 17020 17119: gap of unknown length
 * 17120 22433: contig of 5314 bp in length
 * 22434 22533: gap of unknown length
 * 22534 34645: contig of 12112 bp in length
 * 34646 34745: gap of unknown length
 * 34746 49173: contig of 14428 bp in length
 * 49174 49273: gap of unknown length
 * 49274 67058: contig of 17785 bp in length
 * 67059 67158: gap of unknown length
 * 67159 84928: contig of 17770 bp in length
 * 84929 85028: gap of unknown length
 * 85029 114464: contig of 29436 bp in length
 * 114465 114564: gap of unknown length
 * 114565 165490: contig of 50926 bp in length.

FEATURES
 source
 1.165490
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone_lib="RP11-583F24"
 /clone_lib="RPCT human BAC library 11"

ORIGIN

Query Match 51.3%; Score 4424.6; DB 2; Length 165490;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 5218; Conservative 0; Mismatches 634; Indels 123; Gaps 29;

2713 TTTTAAATCAGACAGCAATTACTTACAGAAAGTTGAACAGATGGACGTGATAGGA 2772
 Db TCTCTCAAACTGCGCACCTGTTTATCAAGCAAGTTGAACAGATGGATGATGGGA 79598
 2773 GAGGCTTAATGATCGATATGATGGACAGAGGCCAAGATCATCTCAAGTTAGATTTGG 2832
 Db GAGGCTTAATGATCGATATGATGGACAGAGGCCAAGATCATCTCAAGTTAGATTTGG 79538
 2833 TCTCAAACTGCTGCGCTCTGATTTGCCATAGTCTCATAGAGAAATACAGACTG 2892
 Db TCTCAAACTGCTGCGCTCTGATTTGCCATAGTCTCATAGAGAAATACAGACTG 79478
 2893 TCCAGATTTGTAAGCCTGATGCTCTCAACAGCTTTCATTGAGCTCTGATGCAATC 2952
 Db TCCAGATTTGTAAGCCTGATGCTCTCAACAGCTTTCATTGAGCTCTGATGCAATC 79418
 2953 TCCGAATTAAGCAACAGAAAGTCTCTGAAAGTCACTGAATCCAGAAAGGCTCTCT 3012
 Db TCCGAATTAAGCAACAGAAAGTCTCTGAAAGTCACTGAATCCAGAAAGGCTCTCT 79358
 3013 ACCTTAGACAGAGGAGGTCTTCAACACTGACCAAGAAAG- GAACGATTAAGGTAAGTA 3071
 Db ACCTTAGACAGAGGAGGTCTTCAACACTGACCAAGAAAGTAAGGTAAGTA 79298
 3072 CCAAGAACTCTCTTCTTCAACAGTCAAGTATGATTTTGTGTAAGATCAATGCTTTATG 3131
 Db CCAAGAACTCTCTTCTTCAACAGTCAAGTATGATTTTGTGTAAGATCAATGCTTTATG 79242
 3132 CTTTCACTTGTGTCATGACAGAGGGGTCACAGAGCTGTTTGAAGAAAGACAGAGAC 3191
 Db CTTTCACTTGTGTCATGACAGAGGGGTCATGAGCTGTTTGAAGAAAGACAGAGAC 79182
 3192 ATGAAGCTTCTTTCAAGAACTGAGTGTCTCAACCAACTGTGTGAGCTTAATGT 3251

79181 ATGAATTTCTTCTGGAACAGTGTGCTGCAACCCAGACTGTGATTTCAACTGGA 79122
 3252 GTCCCCCTTTCAATTTATCTCCCATATCACTCTCTTCAATTCATTCATCTGCC 3311
 Db GTCCCCCTTTCAATTTATCTGCAATATCACTCTCTTCAATTCATTCATCTGCC 79062
 3312 CTCATGGAAGACTGTGCTCTTCAATTCATTTAAGCAAGGAGCATGACAGCAT 3371
 Db CTCATGGAAGAAATGTGCTCTTCAATTCATTTAAGCAAGGAGCATGACAGCAT 79002
 3372 TCTTCCAGATGTAACCTGCTAG- AGCAGTTTCTTGTTCATCTTACTTTCAATT 3430
 Db GCTTCAAGAGTTGAACCTGCTAGAGAGGTTTGTGTTTACTTACTTTCAATT 78942
 3431 TATTTCTTTTGTCTATCTGAAAGGCTTAAGAAAGATATAGATGCCCC-----ATA 3483
 Db TATTTATCTTGTCTGCTGTAAGAGGTTCAAGAAAGAAATAGTGTCCAAACTGAAATG 78882
 3484 ATTAAGAGTGTTCATGAGAAAGTATTAAGAAAGTCAAGAGTTAA----- 3533
 Db ACTGAGAGTGTTCATGAGAAAGTATTAAGAAAGTCAAGAGTTAA----- 78882
 3534 -----GGTCAAGATCTTAAGCAATATAGGGAGCACTACTCTCCCTAG 3585
 Db -----GGTCAAGATCTTAAGCAATATAGGGAGCACTACTCTCCCTAG 78762
 3586 GCTGAAACGACAGGGAAGAGCAGTTCATTTGCGCATATGCCATAGCTGTACCTA 3645
 Db GCTGAAAGGGAAGAGGGAAGAGCAGTTCATTTGCGCATATGCCATAGCTGTACCTA 78702
 3646 AGGCTGAGAGATGAGAGGCAAGTGAAGCCCTGCGGCAAGCAGACCCACA 3705
 Db AGGCTGAGAGATGAGAGGCAAGTGAAGCCCTGCGGCAAGCAGACCCACA 78701
 3706 CAGGCTGATATGTTGATCTGTGTTCCACAAATCTCATGTTGATTTTCCA 3765
 Db CAGGCTGATATGTTGATCTGTGTTCCACAAATCTCATGTTGATTTTCCA 78642
 3766 ATGTTGAAGAAAGGCTGTTGAGAGATATTTAGATCAGGGAGTGTTCATAGA 3825
 Db ATGTTGAAGAAAGGCTGTTGAGAGATATTTAGATCAGGGAGTGTTCATAGA 78581
 3826 ATGTTTAAACATATCCCTTGTGATTTGTTGATGATGACAGATTTCTATGAA 3885
 Db ATGTTTAAACATATCCCTTGTGATTTGTTGATGATGACAGATTTCTATGAA 78524
 3886 TCTAGTTTAAAGAGTGTGACACTCTCTCC--TCTCTTACTCTGCTCACCA 3943
 Db TCTAGTTTAAAGAGTGTGACACTCTCTCC--TCTCTTACTCTGCTCACCA 78467
 3944 TGTGAAGCCTGCTCTCTCTCTCTCTTGTGACAGAGATGGAACCTTCAAGGCTCC 4003
 Db TGTGAAGCCTGCTCTCTCTCTCTCTTGTGACAGAGATGGAACCTTCAAGGCTCC 78407
 4004 CCAGAGCAAGAGCTGATGCTCTTGTGACAGTGTGAGACTTATAGCCAGTTAAAC 4063
 Db CCAGAGCAAGAGCTGATGCTCTTGTGACAGTGTGAGACTTATAGCCAGTTAAAC 78347
 4064 CATTTCTTCAATTAATTTCCAGTCTCAGTTATTTCTTTTGAATTTGAAATGAAT 4123
 Db CATTTCTTCAATTAATTTCCAGTCTCAGTTATTTCTTTTGAATTTGAAATGAAT 78287
 4124 AATACAGACAGAGGACAGAGATGAAATCCCAAGTCTTCTGCTGTCTTCAG 4183
 Db AATACAGAGCAGAGGACAGAGATGAAATCCCAAGTCTTCTGCTGTCTTCAG 78227
 4184 TCTCTGCTGTGTCTCCAGTGTCTCAATTCACAGAAACAGAAATTAAGAAATCC 4243
 Db TCTCTGCTGTGTCTCCAGTGTCTCAATTCACAGAAACAGAAATTAAGAAATCC 78167
 4244 CACTGATGTGTACATAGAGCACTCTTGGAGTGTCAACAGAGATTAAGAAATGG 4303

D	78107	CACGAGTGTGGTGCATAGAAAGCCACTCTCTTGGGGTGTCAAACAGATAGAGAAATGG	78048
Q	4304	AAAGCAATCTCTCAGTGTAAATGAGACTATCCCTCTCACTCTTGTATCTCTTAATTC	4363
D	78047	AAAGCAATCATCATACAGCAATGAGACTATCCCTCTCTCTCTTATATCTCTTAATTC	77988
Q	4364	CTGGGGCTTCTCTATC-----TGATGATCCCTGTCTCATTTGAGCTCTATCAGACTA	4417
D	77987	CTGGGACAGTCTCTGTCTACAGGAGATGATTCCTGTCTATTCAGCTCTTTCACACTA	77928
Q	4418	CTTTAATGTTGGCTGTGCTCTTCTCTACTGTCACTTTATGACGAATGTTGCATTTGT	4477
D	77927	CTTTAATGTTGGGATGTCTTCTCTCTACTGTCACTTTATGACGAATGTTGCATTTCA	77868
Q	4478	TAAAAATGCATAGAAAAATAAAAATGTAATTTTAAAA--AGAACATATGTAATTTGTTTAA	4535
D	77867	CCAAAAATGCATAGAAAAATAAAAATGTAATTTTAAAAAGAGAACATATTTCTTTGTTTAA	77808
Q	4536	ATATTAATGTTGGCTGATCTTATTAATGAACATGAAGAAATATCTTAAACAGAAATAT	4595
D	77807	ATTAACAAGTTGACCCCTCTCATTAAGACAAAGATGAAATATCTTAAATAGAAAGTATC	77748
Q	4596	AGTTGTGCTCTGGGTCATAGGTTCTGAATCTACGATTCACAGTTCACAACTACAGGAGAA	4655
D	77747	AGTGTGTCTCTGGGTCACTAGTCTCTGAATCTCGAATCTCGAATTCACAACTCACAGTGA	77688
Q	4656	CTTTTCCAAAATTAAGGTGTGGCGAGTGTGTATGTACTGACAGGTAACAACCTTGT	4715
D	77687	CTGTTCACAAAATTAAGGATATGGCGAGGTGTGATGTACTGAACAGTATCAAACTTGT	77628
Q	4716	TTTCTTTGTCTATTATTTCTGAAAACTACATATTAACAAGACTTATATGACTTTGCAT	4775
D	77627	TTTCTCTGTCTATCTCTGAAATTAACAAGTAAACAGAACTTACATGCACTTTGCAT	77568
Q	4776	TTTGTGAGTATTTCTTAATTAACCT--TAAATGATTTTAATGATCTGGAGAAAGTGATAG	4834
D	77567	TTTGTGATTAATTTTAAGTAACTTAAGAGATTAATGATCTGGAGAAAGTGATAG	77508
Q	4835	AGTATATATCAAAATPACA-----TATATAGAGAAATGAGCATCTGACGATTTGTGTC	4886
D	77507	GTTTGTGTAACAATCATCATCATTTTATATATAGAAATATGAGCATCTGACGATTTGTG	77448
Q	4887	TGTGCTGGGGTTCTGAAA-----GAAATCCCTGTAAATACAAAA	4928
D	77447	TGTCTGGGGTAGACAGACACATCTGCATATGCGAAAGCCCTGTAAATACAGAA	77388
Q	4929	ATGACACTCTTCAGATCTGAATGAATGAAGAGTCCAAAGCATCATCATCAATGCCAATCCAAA	4988
D	77387	ATGACACTCTTCAGATCTGAATGAAGAGTCCAAAGCATCTGAATGAAGAGTCCAAA	77328
Q	4989	ATTGCTGCTCCCAAGTCTCTAGAGAGTGTGCCATCTCTGTATCTCTACATGATCCCA	5048
D	77327	ACTGCTATCTCCAGTCCCTAGAGAGTGTGCCATCTCTGTATCTCTATATGTGTTCCA	77268
Q	5049	GCGACATTAGCAATCCAGTCTTATAGAAAAAGACAGAGGGAAGAGAGCTTTGTCTCT	5108
D	77267	GCCACACACAGACTTTCAGCATGATGAAAAA--AAATAGAGAAAGAGATGCGTGTCTCT	77209
Q	5109	TCTATTATATCCCATAGACCCAGAGACTTCTCTGTCACTTTTGTGATCTTCCACTTAACA	5168
D	77208	GCTATTATGATCTCTGACCCAGAGACTTGTCTTAATCATCTTTTGTGATCTTCCACTTAACA	77149
Q	5169	GCACTCTGCTACATGGAGATGTCATCCAGACTCAAGAGAAACATGGAGTGGGTCCTGTGCT	5228
D	77148	GCACTCTGCTCATTAATGCAACCCAGACTCAAGAGAAACATGGAGTGGGTCCTGTGCT	77089
Q	5229	GCTTGTACATCTCAGAAAGTTATGTGACAA-----AAAGGAATCTTGGGCAAC	5283
D	77088	GCTTTTAACTCTCAGAAATTTTATGTGACCAACAGAGAAATATGAACATGGGGCAACT	77029
Q	5284	AGCAGTCTCTTACAGCCCTCACTGTCTCTGATTTCTGTGCTCACTCAAGATTTTTCAGGA	5343
D	77028	AGCAGTCTCTTGTGCC---TCTGTCTCTAATCTGTACTCACTCAATCAAGATTTTTCAGGA	76973

QY	5344	ACTCTCAGAAATATATAAAGTGGGGCAGAGAAACGAACTGGAGTCTGTCAGAGATC	5403
Db	76972	ACTCTCTGAAATATATGATTCATCGCCAGAGAAACCATAGAAAGTCTCATGCAAGGCTC	76912
QY	5404	CAGGAGCCAGGGGCTGGTATTTGAGCTGCTCTTCACTGTTGTGAAC-CAGAAAAACCTTT	5462
Db	76912	CAGGAGACAGGGGGCTGGCAITGGGGCTGCTCTTCGTGTTGTGAACCATGAAAAACCTTT	7685
QY	5463	AATTCTCTAGGCTTACTCTTCATCTTATGTTATATGAGATTAATACATATGACAGCTTT	5522
Db	76852	AATTCTCTAGGCTCTGGCTTCATCTTATGTTATATTAAGTAATATACCATAGAGATCTTT	76792
QY	5523	AAAAACATCATATAGCATGTTAAACAACATGCTAAATGTTGGTATACCAACAGTAAAAAG	5582
Db	76792	AAAAAC-----ACATCTAAATGTTGGTATACCAACAGTAAAAAG	7675
QY	5583	ACAGGCATGACTTACTCTTACGAGATCTCGGCTTTCATGAGAGACAAACATATCA-T	5641
Db	76750	ACAGGCATGAC-TAGCGGTTATGAGAGCTTCAGGTTTCATGAGAGACAAACACACATTT	76692
QY	5642	ACCATTAATATAGGAGACAAACGTTTATGTCCTGAGTGGATTAACAGAGGTTCTGC	5701
Db	76691	ACCTTAATATAGGAGACAAATGTTTATGTGCTGCTGCTGATGATTAACAGAGGTTCTGC	76631
QY	5702	TTTTCTCCCATTTCTTTTGGGCAATGAGAGCTGTGGCAGCTGTCTCCCTAAGAGA	5761
Db	76631	TTTTCTCCCATTTCTTTTGGGCAAGAGAGGCTGTGGCAGCTGTCTCTTAAGAGA	76572
QY	5762	GCTCATGATGATGATCACTCACTCTGATGCTCTCTATACTCCAGAGAGATGATCT	5821
Db	76571	GCTCATGATGATGATCACTCACTCTGATGCTCTCTATATCCAGAGAGAGGATGATCT	76512
QY	5822	TCTTTCAACCTGAGAGACTCTGCGCCATGTCATTCGAGATTCGAGAGCAAACTGAC	5881
Db	76511	TCTTTCAACCTGAGAGGCTCTGTCCTGTCATCTCGGATTCAGAGCAAAAGGTGAC	76451
QY	5882	CTCTGATGAGCAAAAAAGAACTCTGATTTGTTCTTAATGAGCAGCACTCACTCAT	5941
Db	76451	CTCTGATGAGCAAAAAAGAACTCTGATTTGTTTAAATGAGCTTCACTCTCTCAT	76392
QY	5942	TTT-TCCCTTATTTGATTTGCTCTCATCTCTCATCGAGGTTGTTTGGTTA---ATT	5997
Db	76391	TTTGTCTCTTATTTGACCTGCTCTCTCTCTCATCTGAAAGTTTGAAGTTAGTTT	76331
QY	5998	TTTTTTTTGAGCCCAATTTTGACTGTCACTTGATTTTAACTTGAGATCACTCTCT	6057
Db	76331	TTTTTTTTGAAACCAACAAATTTTGCGGCAACTTGATTTTATCTTGAGATCCCTCTCT	76272
QY	6058	ACTTTACCCCCCTTAAACATGTTATATCCACACATATGTTGTTGGTCCAAAGGCTGG	6117
Db	76271	CTTTTACCCG-CTCTTAACATGTTATATTCGACACAGAGTGTGTGGTGTAAAGGCTGC	76213
QY	6118	TGAAAAATGGAATCATAGTCAGAGCCGCTGCGGCTCACAATTCATATATATATATA	6177
Db	76212	TCAAAAGCTGATCATATAGTCAGAGCCGCTGCGGCTCAATTCATAC-----TATATA	76162
QY	6178	CCCCCGGACAAATATATATCTCTCTTTTACTCTAATTTCAATATCTCAATACAGAA	6237
Db	76159	GTCACCAATGAAATTAATTCACCTCTTTTGTGTTCCAA-TGCATTTATCTACATATGGAA	76101
QY	6238	TAAATCAATATTTTAACTCTAGAGCTTCAAGATGATTAAGAGGCAATPACTATTA	6297
Db	76101	TAAATCAATATTTTAACTCTAGAGCTTCAAGATGATTAAGAGGCAATPACTATTA	76041
QY	6298	CTGTCAATACAGCTGCTGTTATTTCTCCCAATTAAGACTATCTCTATCTCCAGTTGAA	6357
Db	76040	CTGTCAATACAGCTGCTCTTATTTCTCCCAATTAAGACCAAGTCTCATTTCTCCAGTTGAA	75981
QY	6358	TTTGCATGAATATCTCTTTTAAACCAAGGCTTCACTTCTCTATTTTCACTCATAG	6417
Db	75980	TTTGCATGAATATCTCTTTTAAATCTCAAGGCTTCAATATCTCTATTTTGGCTCATAG	75921

OY	6418	ACTCCCTCATCAAAATGGTTTGCATCAAAAGAAAGGCTACAAAGATCTCCGAAAGA	6477
Db	75920	ACTCCCTCACTACAAATGGTTTGCATCAAAAGAAAGGCTACAAAGATCTCCGAAAGA	75862
OY	6478	GAGAAATGAATAGGTTTACATTTGGTATCTCAGCAGAAACTTAGTAGTCCCCATCA	6537
Db	75861	AGGAATGAATAGGTTTATATTTGGTATCTCAGCAGAAATCTAGTAGTCCCCACGA	75802
OY	6538	TATTTCCACACTTCAATTTACCTGCTGCAAGTGGCACTCAGGCTCACCTCCTTACTT	6597
Db	75801	TATTTCCACACTTAAATTTACCTGCTGCAAGTGGCACTCAGGCTCACCTCCTTACTT	75742
OY	6598	CCTCGTCTATTTGCTGAGCAATTTCACTCAGCAACCAACCTTACC-AAACATGTGA	6656
Db	75741	CCTCGTCTATTTGCTGAGCAATTTCACTCAGCAACCAACCTTACCACATGTGA	75682
OY	6657	CAAAATGCTTCTAGGGGTTGGCAAGCACTCAGTGCCTTATTTTAAAGCATCAG	6716
Db	75681	CAAAATGCTTCTAGGGGTTGGCAAGCACTCAGTGCCTTATTTTAAAGCATCAG	75622
OY	6717	TGCTCAATTTCAAGTTTGGGCACTCAATCAATCTTCTCAACACAGATAGACTGTC	6776
Db	75621	TGCTCAATTTCAAGTTTGGGCACTCAATCAATCTTCTCAACACAGATAGACTGTC	75582
OY	6777	CACAAATGAATTTCTGATGAATGAATTTTCTTATCTAATATATGTGTGTTCTAT	6836
Db	75561	CACAAATGAATTTCTGATGAATGAATTTTCTTATCTAATATATGTGTGTTCTAT	75502
OY	6837	GCTTCAATGTGCTTTCATTTTATTTTCCATTTCAATCCAAATATACCATTCGCATTAG	6896
Db	75501	AACTTCAATGTGCTTTCATTTTATTTTCCATTTCAATCCAAATATACCATTCGATTAG	75442
OY	6897	GCTTCTCATGATGATTCCTTCATTTGATGAAGGTTATGAAGAATTTGCTGCT	6956
Db	75441	GCTTCTCATGATGATTCATTTCAATTTGATGAATTTATGAAGAAGCAAGTGTCTGCT	75382
OY	6957	TATGAATAGGCACTAGAGAGTAAATATGTAAATATGTGCTCTGTCTGCAATGATGACA	7016
Db	75381	TATGAATAGGCACTAGAGAGTAAATATGTAAATATGTGCTCTGTCTGCAATGATGACA	75322
OY	7017	CAGTGAATTTTCTCAACCCACAGAGTCCGSCATTTTCAACAC-ATCTAGCAAGATC	7074
Db	75321	TGCTGAGTTATTTCTCAACCCACAGAGTCCGSCATTTTCAACATTTCTTAGCAAGGTC	75282
OY	7075	CCATTTTCTCTGAGTTCAATATGATGATCTTTTCTCTGTCCAGAGATGACAGTCTCG	7134
Db	75261	AAATTTTCTCTGAGTTCAATATGATGATGCTCTTCTCTGTCCAGAGATGACAGTCTCG	75202
OY	7135	GTATGAGGGGTGTCAACAACCTCTTTGTGTATTTGTAATTTCTCCTCAGAGTGAAGAAAT	7194
Db	75201	GTATGAGGGGTGTCAACAACCTCTTTGTGTATTTGTAATTTCTCCTCAGAGTGAAGAAAT	75142
OY	7195	TTTCAGGCCCAAGATAGAGTATCATCGGGTCAACAGCACTGGCTGATGAGGGGGGT	7254
Db	75141	TTTCAGGCCCAAGATAGAGTATCATCGGGTCAACAGCACTGGCTGATGAGGGGGGT	75082
OY	7255	TTTGTATCTTATGTTATCCCATGTCAAGACAGAACTTGTGTGCGAGTAGAGAGGTCA	7314
Db	75081	TTTGTATCTTATGTTATTTCCATGTGTGACAGAACTTGTGTGCGAGTAGAGAGGTCA	75022
OY	7315	GGCTTACAGATCAACAGAACTGGAATTTCAAACTGGAATTTAGAACCCCACTTTTGTAT	7374
Db	75021	GGCTTACAGATCAACAGAACTGGAATTTCAAACTGGAATTTAGAGAACCCCACTTTTGTAT	74962
OY	7375	AGGTGACTTATCTCTGAGGCTCTGATCTCTCTCTTAAATAGGACATTAATCCC	7434
Db	74961	AGGTGACTTATCTCTGAGGCTCTGATCTCTCTCTTAAATAGGACATTAATCCC	74903
OY	7435	ACATGCGAGGGTGTGGGAGAAATCAGATCAAAACACTGGTGTATCATCTGTTTCT	7494
Db	74902	ACATGCGAGGGTGTGGGAGAAATCAGATCATACACTGGTGTATCATCATGTTTCT	74842
OY	7495	GTTTCCAGGGTATCAGACTGGGGTTCCTGACATGATTCACCAATCCCACTTGGGT	7554

Db	74442	GTTTCCAGGGTCAACGACTAGGGTTTCTGAGATGAGATCCAACTCAACTTCAACTTTGGAC	74783
QY	7555	ACAGAACTGAACACCAATCAACGAGCTGAGAGACTCTTGTCTAACAGCAGACCTTGAGC	7614
Db	74782	ACAGAACTGACACCAATCAACGAGAACTGAGAGAGACTTTGTCTAACAGCAGACTTTGAGC	74723
QY	7615	TTTACGGGGGTGAGAGTGCATTCGTTTTCCCTTTGTCGGCTGACAGGAAAGCGGGTTGTGCTC	7674
Db	74722	CTACAGGTGTGACAGTGCATTCGTTTTCCCTTTGTGCGGCTACAGGAAAGCGAACTTGTGCTC	74663
QY	7675	TGGCTCTGTGGGCTGCGGCATGCGCAGAAAGCGTGTCTCATCTAGACTCTCAACTGTGTC	7734
Db	74662	TGGCTCTGTGGGCTGCGGCATGCGCAGAAAGCGCTTCTCATCTAGACTCTCAACTTGTGCTC	74603
QY	7735	GCGGCGCACTTCTCTTCTTTAGCGGCGCAATTATATGTTGCGCGTTACGCTCATCAAT	7794
Db	74602	GCAGCAGACTTCTCTTCTCTCTCAGCGGCGCGCTTATATATTTCCCTGTAAAGCTTCACTAGT	74543
QY	7795	ATCCGCACTCCCATTTTCGAAAAATTCCTCAAGTCCCTGTGTGATGACCTTTCCCTATTATAGC	7854
Db	74542	ATCCCCCATCACTCTCTTAAATCCCTCTATCCCTGTGATGATGTCTTCTACTTTGACAGC	74483
QY	7855	CTAAGCATGTGTGAGCGGCATCAGACACGAGCGCTGCTGTCCATCCCTGTGGCCCATCTGG	7914
Db	74482	CTGAGCTTTCTGAGTCCGCTGAGCAGCAGAGCGTGTCTGTCCGTCTGTGGCCCATCTGG	74423
QY	7915	TACCACTGCGCGCGCGCCAGATACCTGTCACTGGGTATGTGTCTGTCTGTGGGCCCTGG	7974
Db	74422	TACCGCTGCCACCGGCCCAACACACTGTGACGGGTGTGTGTGTCTGTCTGTGGGCCCTGG	74363
QY	7975	TCCCTGCTCGGAGATATCTGTGAGTGTGATTTCTGTGACTTCTCTGTTTATGTGTGAT	8034
Db	74362	TCCCTGTGTGGAGCATCCCTGAGTGTGATGTATGTGTGCTTCTCTGTCACTGTGTCTAT	74303
QY	8035	TCTGTTTGTGTGAAAGTACAGATTTCAATTACATGCGGTGCGGGTTTTTTTATGTGTG	8094
Db	74302	TCTGCTTGTGTGTCAACATCAGATTTCAATCACTGAGTGTGCTGTGATTTTTTTATGTGTG	74283
QY	8095	GTTCCTGTGTGGGTCCAGCTGTGCTGTGTGTGAGGATTTCTGTGTGATCCCGGAAAGTGT	8154
Db	74242	GTTCCTGTGTGGGTCCAGCTGTGCTGTGTGTGATCAGGATTTCTGTGTGATCCCGGAAAGTAT	74183
QY	8155	CCGCTGACAGAGCTGTGATGACCAATCTCTTCACTGAGTGTGTCTTCTCTCTGTGGC	8214
Db	74182	CCGCTGACAGAGCTGTGATGACCAATCTCTGTCTCACTGATGTGTCTTCTCTCTGTGGC	74123
QY	8215	CTGCGCTTTGGCATTCAGTGTGGGCGCTTTTCCAGGATCAACTGGATTGGAAAGTCTTA	8274
Db	74122	CTGCGCTTTGGCATTCAGTGTGGGCGCTTTTCTATATTTATGTGATCCAGTGGACAGGAAAGTCTTA	74063
QY	8275	TTTTGTGATGTGATCTAGTTTCCATTTTCTGTCCGCTCTTAAACAGCAGTCCAAACCC	8334
Db	74062	TTTTGTGATGTGATCTAGTTTCTATTTTCTGTCCGCTCTTAAACAGCAGTCCAAACCC	74003
QY	8335	ATCATTTACTTCTTCTGTGGGCTCTTTAGGCAAGCTCAAAATATGCGAACTTGAAGCTGTG	8394
Db	74002	ATCATTTACTTCTTCTGTGGGCTCTTTAGGCAAGCTCTCAAAATATGCGAACTTGAAGCTGTG	73943
QY	8395	GTTCCTCAAGAGGCTGTGACGAGACAGCGCTGTGAAGTGTGATGAAAGTGTGAGAGGAGCTTCT	8454
Db	73942	GTTCCTCAAGAGGCTGTGACGAGACAGCGCTGTGAAGTGTGATGAAAGTGTGAGAGGAGCTTCT	73883
QY	8455	CAGAAAACCTGTGAGCTGTGCGGAAACAGATTTGAGACATGTGAGGAAAGAACTTGTCCCTG	8514
Db	73882	GAGAAATCTGTGAGAGCTGTGCGGAAACAGATTTGAGACATGTGAGGAAAGAACTTGTCCCTG	73823
QY	8515	TCAGACAGAACTTTGAGAGCAATGTGCGCTGTGCACCTTGAACAATATATGATGATTTTC	8574
Db	73822	TCAGACAGAACTTTGAGAGCAACGTGCGCTGTGCACCTTGAACAATATATGATGATTTTC	73763
QY	8575	TTAGCTTTGTGCTCAGAAATGTCTCAGGGTCCCC 8609	

Db 73762 TTAGCCTTCCTGACAGAAATGCTCAGTGTCC 73728

RESULT 9
AC103974
LOCUS AC103974 189230 bp DNA linear PRI 13-MAY-2002
DEFINITION Homo sapiens chromosome 11, clone RP11-1081L13, complete sequence.
ACCESSION AC103974
VERSION AC103974.6 GI:20531887
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-1081L13
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamas, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL
REFERENCE 3 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamas, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,

TITLE
JOURNAL
COMMENT
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamas, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 13, 2002 this sequence version replaced gi:20330974.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21856
Center clone name: 1081_L13

FEATURES
source
Location/Qualifiers
1..189230
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-1081L13"
/clone_1bp="RP11-1081L13 Human Male BAC"
198..223
/rpt_family="CTCA n"
1203..1215
/note="<30 qual SNGL region"
2295..2378
/rpt_family="L2"
complement(2381..3394)
/rpt_family="L1MD2"
3396..4062
/rpt_family="L2"
4172..4245
/rpt_family="AT_rich"
4411..4465
/rpt_family="TA n"
complement(4482..4546)
/rpt_family="L1MC4a"
complement(4547..4857)
/rpt_family="AluJo"
complement(4858..5400)
/rpt_family="L1MC4a"
6131..6239
/rpt_family="L2"
complement(6654..7008)
/rpt_family="WL71B"
7551..7749
/rpt_family="L2"
8025..8057
/rpt_family="AT_rich"

	repeat_region	9218 . .9800	/rpt_family="L1PB1"	
	repeat_region	9883 . .9955	/rpt_family="L1PB1"	
	repeat_region		/complement(11450 . .12969)	
	repeat_region		/rpt_family="L1PA4"	
	repeat_region		/complement(13902 . .14364)	
	repeat_region		/rpt_family="L1R33"	
	repeat_region		/complement(14660 . .15009)	
	repeat_region		/rpt_family="MER2"	
	repeat_region		/complement(15115 . .15369)	
	repeat_region		/rpt_family="MLT1J"	
	repeat_region		15517 . .15603	
	repeat_region		/rpt_family="MIR3"	
	repeat_region		/complement(15612 . .15692)	
	repeat_region		/rpt_family="L2"	
	repeat_region		16230 . .16337	
	repeat_region		/rpt_family="MIR"	
	repeat_region		/complement(16959 . .17034)	
	repeat_region		/rpt_family="L2"	
	repeat_region		17340 . .17464	
	repeat_region		/rpt_family="MIR"	
	repeat_region		19511 . .19704	
	repeat_region		/rpt_family="L1MB5"	
	repeat_region		19705 . .19813	
	repeat_region		/rpt_family="L1MB4"	
	repeat_region		20596 . .20683	
	repeat_region		/rpt_family="L2"	
	repeat_region		/complement(20686 . .20691)	
	repeat_region		/note=<30 qual SNGI region"	
	repeat_region		/complement(20691 . .20874)	
	repeat_region		/rpt_family="L1MD3"	
	repeat_region		21261 . .21416	
	repeat_region		/rpt_family="MIR"	
	repeat_region		/complement(22006 . .22307)	
	repeat_region		/rpt_family="L1BP4"	
	repeat_region		/complement(22308 . .22559)	
	repeat_region		/rpt_family="Alusg1"	
	repeat_region		/complement(22560 . .24537)	
	repeat_region		/rpt_family="L1P4"	
	repeat_region		24532 . .24718	
	repeat_region		/rpt_family="L1"	
	repeat_region		/complement(24717 . .25049)	
	repeat_region		/rpt_family="L1P4"	
	repeat_region		/complement(25039 . .25908)	
	repeat_region		/rpt_family="L1PA4"	
	repeat_region		/complement(25885 . .27145)	
	repeat_region		/rpt_family="L1"	
	repeat_region		/complement(27179 . .29623)	
	repeat_region		/rpt_family="L1BP4"	
	repeat_region		/complement(29632 . .30526)	
	repeat_region		/rpt_family="L1"	
	repeat_region		/complement(30530 . .31245)	
	repeat_region		/rpt_family="L1MBG"	
	repeat_region		/complement(31246 . .31555)	
	repeat_region		/rpt_family="Alusx"	
	repeat_region		/complement(31556 . .31800)	
	repeat_region		/rpt_family="L1MBG"	
	repeat_region		/complement(31853 . .32053)	
	Query Match	30.2%; Score 2601.8; DB 9; Length 189230;		
	Best Local Similarity	72.3%; Pred. No. 0;		
	Matches 4441; Conservative 0; Mismatches 1072; Indels 626; Gaps 57;			
OY	2591 CAAAAATTGTGTTCAGGCCATTTGCAGCCTTGAAATGATGCCATGACTTAATTTTCATT			2650
Db	13023 CACAGAGTGCGATCAACCCATTTGCAAGCCCTGAAGAACCCTATGATCTAATTAGTTTT			130822
OY	2651 GGATATAAGCCCTCATATGATATATTCACATTTGCTT -AATGGATGCCATAGCCCCATAT			2709
Db	13083 AGTTATTAAGCCCTCACAAAACAAGCTCAGTTGCTTGACATTTGATTAACCTTAGGCCATTAAT			131422
OY	2710 TGAATTTTAAA-ATCAGAGACGCAATTAATTACAGAAAGTTGAMCAAGATGGAGCGTAT			2768

Db	13143	AAATGTAAAAAGTTCAAACACGAAATTACTTACAGAAAGTTGAATGAATGGATATGAT	13202
Qy	2769	AGAGAGGCTTAAATGTACTGTGATATGGACAGAGGCCAAGAAATCTCAGTTAGATT	2828
Db	13203	GAGAGAGGACAGATATGTATTAAATATGGGACAGAGGCCAAAGAAACATCTCAGTTAGGGAT	13262
Qy	2829	TGTGTCTCAAAATACCTGTGGGCTCTGATTTGGCCATATGTCCTCATACAGAAATTAACAG	2888
Db	13263	CGTGCTCTCAACAGCTGTGGCTCTGTATGTTGGCTGTGCAATCTCATACAGAAATCACATG	13322
Qy	2889	ACTGTCCAGCATCTTCGTAAAGCCTGTGATTGTCTACAGAGCTTTCATTTCAGCTCTGTAG	2948
Db	13323	--CTGTCTGATATCTTAAACCTGGGTTGGTTGCCAGTTCTCATTCAGCTCTGTAGG	13380
Qy	2949	CATCTCCTGAATTAAGCAACACAGAAAATGCTCTGTAGTCACTGAATCCAGAAAAGCT	3008
Db	13381	CATCTCCTAAATGACACACACAGAAATAGTCTGTGAATCAACGGAATCCAGAAAAGCT	13440
Qy	3009	CTCAACCTTTAGCACAAGGAGGCTTACG-----ACTGGAACAAGAAAGCAATAG	3063
Db	13441	CTACCCCTTTAGCAGAAGGAGGAGCTTTATCTTTTGACTTGAAGAAAGAGAAAGTTAG	13500
Qy	3064	GGTAGTAGCAAGAACTCTCTTC--TTCACAGTCAATTATGATTTTGTCTGAATCAT	3122
Db	13501	GGTATGTATCCAAAGAACTCTTCTTCTTCCACAGAAATATGTGTTTGTGTATTAAGTTAC	13560
Qy	3123	GTCCTTATGCTTCACCTT--GGTGCTACATGCAAGGGGGTACAGAGCTTGTTCAGAAA	3180
Db	13561	AAACCTGTGTTTCCACCTTGGGGTGTGCAATACAGAAATTCATTAATTTCTTCAAGAAA	13620
Qy	3181	AGACAGAGACATGAACCTCTTTCAGAACTGAGGCGTCAACCCAACTGTGTGAG	3240
Db	13621	AGCAGAGACATGAACCTTCTTCTTGAACCTGAGTGTGCGACGCCAAGCTGGGGAG	13680
Qy	3241	CTCTAAATGTGTCTC-----CCCCCTTCAATTTATCTCCCATATCACTCTCTTC	3290
Db	13681	TTTCAAAATGGGGTCAGACAGCTTTCCTCTGTATCAATCTCTCATATCACCCCTTT	13740
Qy	3291	ATTCCAAATCATTCATCTGCCCTCATGAGAGACGTGCTCTTACATTCATTTAACGA	3350
Db	13741	ATTCCAAATCATTCATCTGCTTTCATGAGATATATATGCTTTTAAATTCATTTAAGGA	13800
Qy	3351	GCAAGGGGACATGACGCAATTTCTTCCACAGTTGAATGCTATTAAGCAGATTTCTTGG	3410
Db	13801	GCAAGGGGATATGACGCAATCTTCTTCTACAGCTGAGCTGTGTTAAG--GGCTTATTA	13855
Qy	3411	TTTCACTTACTTTTCAAAATTTATTTCTTTTGCCTATCTTGAAAGGCTTAAAGAAATAT	3470
Db	13860	TTTTACTTACTTTTCAAAATTTATTTCTTTTGCCTATCTTGAAAGGCTTAAAGAAATAC	13919
Qy	3471	AGATGGCCCC-----AATAATTAAAGAGTTCATGAGAAAGATTTACAAANATGC	3523
Db	13920	AGATGTCTCAAAATGAATTAATGAAGAGTTCATGAGGGGACATATGACAAATATG	13979
Qy	3524	ACAGAGTTAAGG-----TCAGATCTCTAAGACACAATATACATAGGGAGCA	3569
Db	13980	ACAGGTTTAAGGGAAAAAGACAATGCTTAAGGAAGACCTTAATATGGCAATGTGTAGCA	14039
Qy	3570	CTACTTCTCTCCCTAGGCTGAACGACAGGAAAGACGATTACATTTG-----TGGC	3623
Db	14040	CTACTTCTCTCCCTAAGCTTAAAGGACAGGAAAGAGCATTAACTTGAAGACACAGC	14099
Qy	3624	CATAGCCATAGCTGTAGCATTAAGGGTGGAGACATAGACAGGCAATGAGAAAGCCCT	3683
Db	14100	CATAGCCATAGCTGTAGCTGTAGGGTTGGAGAGATATGACAGGACGGTGGAGAAACCTT	14155
Qy	3684	GGGTGGCCAAAGCAGCAGACACACAGGCTGTATATAGTTGGATCTGTGTTCCACCAAAT	3743
Db	14160	GCATACCAACACACAGCCACAAAG-----	14188
Qy	3744	CTCATGTTGATGTATTTCCAAATGTGAGAAAGGCTTGTGGAGATGATTAATTAAG	3803

Db 14186 ----- 14185
 Qy 3804 TCACGGGAGTGTTCATGATGATTTTAAACACCATCCCTTGTGATTTGTTGTG 3863
 Db 14186 ----- 14185
 Qy 3864 ATACTGACGAGTTCATGAAATCTAGTTTAAAGCGTGTAGCACTCTCTCTC 3923
 Db 14186 ----- 14185
 Qy 3924 TCTTACTCTGCTCTCACCATGTGAGACGCTCGCTCCCTTGTGCTTTTACACAGAT 3983
 Db 14186 ----- 14185
 Qy 3984 GGAAGCTTCTGAGGCTCCCGAAGAGAGCTGATGCTTCTTGTGACATCTGTAG 4043
 Db 14186 ----- 14185
 Qy 4044 AGCTATTAGCCAGTTAAACCATTTCTTCATTAATTTCCAGTCTCAGTATTTCTTT 4103
 Db 14186 ----- 14185
 Qy 4104 TAGCAATTTGAGATGAACTAATACACAGAGAGCCAGAGATGGAATCCCAAGT 4163
 Db 14186 ----- 14215
 Qy 4164 GCTTTCCTGCTGTCTTCAGTCTCTGCTGTGCTCTCCAGTGTCTCAATTCACAGAA 4223
 Db 14216 ACTGCTCTGCTGCTCTCAGTCTCTACAGTGTCTCAGTGTCTCAGTCTCAGTCTCAG 14275
 Qy 4224 ACCGAAATTAAGAAATCCCATGTATGTGTATATAGAACCACTCTCTGGATGTCA 4283
 Db 14276 ACCA-AAAGAAATATCACTGATGATGATGACAGAACCCCTCTGGATGTCA 14334
 Qy 4284 AACAGATTAAGAAATGAGAAAGCAATCTCATGTATATGAGATCATCTCTCTCACC 4343
 Db 14335 AACAGGATGAGAAATTTGGAGAGCAATCTCTCAAGCAATGAAATCTCTCCTCACC 14394
 Qy 4344 TTTCTGTA--TCTCTTAATTTCTGAGGCTTTCTCT--ATCTGATGATCCTCTG 4394
 Db 14395 TTTTATATCTTCTCTCCTACCTCTCGGGTCTTCTCTCTAATCTCATGATCATGT 14454
 Qy 4395 CTCATTTCACTCATGACA-CTACTTAATGTTGTGCTTCTCTCTCTCTCTCTCT 4453
 Db 14455 CTCATTTAATCTCTATGAGATTTTATTTTATGTTGAAATCTTCTCTATGTCGAT 14514
 Qy 4454 TTATGCAAAATGTTGATTTGTTAAATGATGATGAAATTAATTAATTTTAAATA 4513
 Db 14515 TTATGAGAAATCTCTGATTTATCAAAATGCTCTAGAGATGAAATTTTAAATA 14574
 Qy 4514 G---AACATATGATTTTGTGTAATATAGTTGTGCTGATCTAATTAAGACATGAG 4569
 Db 14575 GAGAAACATATCTCTTTTGTGTAATGATTAAGTTGTGCTCTCTAATAAGAACAGAG 14634
 Qy 4570 AAGAAATCTTAAACAAAGATATAGTTGTGCTCTGAGTCACTAGTCTGATCTA 4629
 Db 14635 TAGAAATATCTTAAACAAAGATGACAGTTGTCTCTCATGTCACTGTGCTCTGATC 14694
 Qy 4630 CAGATTCAAACAACTACAGAGAGAACTTTTCCAAAATA-AAGTGTGAGGAGGTTGTG 4688
 Db 14695 TTAGTTCAACAAACACAGATGATATATTTCCAAAAAAGAGAGACAGGATGCTGTA 14754
 Qy 4689 TATGTACTGAACAGATCAAACTGTATTTCTTTGTCTATTAATTTCTGAAAACTACA 4748
 Db 14755 TAGGTACTGAACAGATCAAACTTTCTTTCTTCTATTAATTTCTTAAACATACATA 14813
 Qy 4749 TAAAGAACTTATATGATTTGATTTGTGATTTCTTAATA-CTTAAATGAT 4807
 Db 14814 TAAACATTAATTAATGATCTTCAATTTTAAAGATTTATTAAGTATCTAGAAATAT 14873
 Qy 4808 TTAATGTATCTGAGAGAAAGTATGATATATCAATACCA-----TATATAG 4859
 Db 14874 TTAATGTATTAAGAGAAAGTTCATAGTATATTTCAAAATCTACATCTTTCTATAG 14933

Qy 4860 GAAATGAGATCTGACAGATTTTGTCTGTGCTGAGGCTTCTGAAAGAAATCCCTGTA 4919
 Db 14934 GAAATGAGATCTGACAGATTTTGTGTGTGCTGAGGCTTCTGAAAGAAATCCCTGTA 14993
 Qy 4920 TACACAAAAAGACATCTTGAATCTGAATCTGAATCTGAATCTGAATCTGAATCTGA 4979
 Db 14994 TACTGAGAGATCTGATTTTAAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 15053
 Qy 4980 ATTCCAAAAATTTGCTGCTCCCACTTCTAGAGATGCTGCTCTGATCTGATCTGATCA 5039
 Db 15054 ACTCCAAAAATTTGCTGCTCCCACTTCTAGAGATGCTGCTCTGATCTGATCTGATCA 15112
 Qy 5040 TGTCTCCAGGACATTAAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 5099
 Db 15113 TGTCTCCAGGACATTAAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 15171
 Qy 5100 TTTGCTCTCTATTAATCCCATGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5159
 Db 15172 TGTCTCTCTATTAATCCCATGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 15231
 Qy 5160 CACTTAAACAGACCTGCTCATGAGATGATCTGATCTGATCTGATCTGATCTGATCTG 5219
 Db 15232 CATGACCAAGACCTGCTCATGATGATCTGATCTGATCTGATCTGATCTGATCTGAT 15291
 Qy 5220 CTTGTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 5275
 Db 15292 CTTGTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 15351
 Qy 5276 GGGACACAGAGCTCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 5335
 Db 15352 GGCATGAGATGAGCTCTGAAAGCTCTGATATATCTGATATCTGATATCTGATATCT 15411
 Qy 5336 TTTGAGAACTCTCTGAGAAATTAATTAATGATGAGGAGAGAAAGAACTGAGATCT 5395
 Db 15412 TTTGAGAGATCTCTGAGAAATTAATTAATGATGAGGAGAGAAAGAAAGTCTCAT 15471
 Qy 5396 CAGATCTCAGAGACAGAGGCTGATTTGAGACCTGCTCTGATGTTGTGAAC-CAGAA 5454
 Db 15472 CAGGCTCTCAGAGACAGAGGCTTATCTGATCTGCTCTGATGTTGTGAAC-CAGAA 15531
 Qy 5455 AACCTTAAATCTCTGAGCTTATGCTTATGATGATGATGATGATGATGATGATGAT 5514
 Db 15532 GATCTTTGATCTTTTAAAGCTCTGAGAGCTTCTGATCTGATGATGATGATGATGAT 15591
 Qy 5515 CAGCTTTTAAAGACATCATGATGATGATGATGATGATGATGATGATGATGATGATG 5574
 Db 15592 TGTCTTTGAGAAC-----ACAGGCTTAAATGCTGCAATCTACAG 15633
 Qy 5575 TGAAGAG-ACAGGATGATCTTACTCTTACGATCTTGGGTTTATGAGAAAGACAA 5633
 Db 15634 TGAACAGAGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 15693
 Qy 5634 CATATCATCATCATCATATGATGATGATGATGATGATGATGATGATGATGATGATG 5687
 Db 15694 TATATCATCATCATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 15753
 Qy 5688 AACAGAGTCTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5747
 Db 15754 AACAGAGTCTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 15813
 Qy 5748 GTCTCTTAAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5807
 Db 15814 GTATCTCTTAAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 15873
 Qy 5808 AGAGAGATGATCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5867
 Db 15874 AGAGAGATGATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 15933
 Qy 5868 AGAGCAACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5927
 Db 15934 AGAGCAACGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 15993

QY 5928 GCACTACCTTATTT-TCCTTATTTGCTTCTCATCTGATCGAGTTGT 5986
 DB 15994 TCACCTACCTATTTGGTTCCTATTTGATTTGCTTCTCATCTGATCGAGTTGT 16053
 QY 5987 TTAGGTTA-ATTTTTTTTTTACGCCACAATTTGACTGCACTTGATTTAACTTGA 6045
 DB 16054 TTAGGTAGCTTTTTTTTTTACGTCMAACTTTGACTGCACTTGATTTAACTTGA 16113
 QY 6046 AATGCTCTCTACTTATACCCCTCTTACATGATATATGACATAGTGTGTGGT 6105
 DB 16114 AGTCATTTCTGTGCTT-CCCCCTCATCATATATATGACATAGTGTGTGGT 16171
 QY 6106 CCAAGGGCTGTGAAAA-ATGATCATGATGACGCTGTGGGCTCATCTCATCT 6164
 DB 16172 CCAAGGGCTGTGAAAAAGGTGACCTTGAGGACGCTGTGGGCTCAAACTCCGACT 16231
 QY 6165 ATATATATATATACCCCGGACAATATATCTCTTTATATCTTAACTTCAATTC 6224
 DB 16232 GTGTAAAG-CCAGGGACAATATATCTCATCTTTTGTGCTTAACTTCAATTC 16284
 QY 6225 TGCAATACGAAATATATATTTTAACTCTCTAGGCTCTTCAATGATTTAAAGGC 6284
 DB 16285 TGCCATATAGAAATATATCTCATCTCTAGGCTCTTGAAGCATTA---AGAC 16340
 QY 6285 AATACCTATTAATCTGTCATGCTGTGTATTTCCCAATTAAGCTTAATCTCAT 6344
 DB 16341 AATCCCTATTAATCTGTCATGCTGTGTATTTTCAAGCAATATAGATTAATCTCTGT 16400
 QY 6345 TCTCGATTTGAATTTGATGAATATCTCTTTTAAACCAAGCCCTTACATTTCTCTA 6404
 DB 16401 TCACCACTGTAATCTCATGATGATCATCTTTTTCATTCMAAGCCGTGAA---CTCTTA 16457
 QY 6405 TTTCCATGATGACCTCTCTCATACAAATGTTGATCAACAAAGAAAGCTTACCAAG 6464
 DB 16458 TTTCTCTATGACCTCTCTCATACAAATGTTGATCAACAAAGAAAGCTTACCAAG 16517
 QY 6465 ATCTCCGAAAGAGAAATGAATATGATTTTCAATTTGATGATCTGACAGAACTTATG 6524
 DB 16518 ATCTCTGATGAGACAAATCAATGATTTTGTGACCTGTGACACCAAGAAACATTTAG 16577
 QY 6525 AGTCCCAATCATATTTCCCACTTCAATTTACCTGTGACGTGACCTGAG- GCTTACC 6583
 DB 16578 AGTCTCCCAAGATATGCTATCTTCAATTTACCTTTTCAAGGCGCTCATGACCTCAC 16637
 QY 6584 CTCACCTACTCTTCTCTGTTCTATTTGAGCAATTTGACCTCAAGCCCAACCCCTACC 6643
 DB 16638 CAGGCTTACTCTTCTTCTTCTATATTTGAGCAATTTGAGCTGAG- CCAACCCCTACC 16696
 QY 6644 CAACACTGTGTACAAA- TGCTTCTAGAGGTTCCGCAAGCACTGAGCTTATTT 6702
 DB 16697 CCAACACTGTGTACAAAAGATTTCTAGAGGCTCAGAAAGCCACTGAGCTCTCATTT 16756
 QY 6703 TAAAGGACATGATGTGCAATTTCAAGT- TTTGGGCACTCATCATCTTCTTCAAC 6761
 DB 16757 CAAGGACATGATGTGCAATTTCAAGTATTTGGGCACTCATCATGATTAAGCTTTTAA 16816
 QY 6762 ACAGATAGAGCTGTCAAAATAGATTTGATGAATGAATTTCTCTCAATTTATA 6821
 DB 16817 ACAGAGGA-ACAAATATATTTGATGAATGAATTTT- TACATCTCATTTATA 16868
 QY 6822 TGTGTGTGTCTTAATGCTTATCATGCTTCTTCAATTTTATTTTCAATTTCAATCAATC 6881
 DB 16869 TGTGTGTGTCTTAATGCTTATGATTTTATTTTATTTTCTTTTAAATCAAGC 16928
 QY 6882 TACCATTTGATTTAGGCTTCTCATGATGATTTCTTCAATTTGATGAATGAATTTATA 6941
 DB 16939 TATCATCTATCATTTAGGCTTCTCATGATGATTTTCACTGATGATTTTATACAAA 16988
 QY 6942 GCACATTTGCTGCTTATGGA- ATAGGCACTAGAGATTAATTAATTAATTTGCTGCTG 7000
 DB 16989 ACACATTTGCTGCTATGAGCTTAGGCACTGAGGATATATAA-----GTGCTCTC 17040
 QY 7001 TCTGCAATGACTG-ACACACTGATTTATTTCTACCAACAGGTCCCGCATTTTCA 7058

DB 17041 TCTGTATATAGGAGACACACTGAGCATTTCTTACTACCAAGGCCAGTATATCA 17100
 QY 7059 CAT---CTTAGGAAATCCATTTTCTCTGTGCTATATGATGATTTTCTCTGT 7115
 DB 17101 CTGTCTCTAGCAAGAGCTCATATTTTCTTCAATGATTAATTTTCTTACCTTATTTGAT 17160
 QY 7116 CCAGATATAGCAGTCTGTGATGAGGTGTCACAAACCACTTTTGTATGATTAAT 7175
 DB 17161 CCA-TGATGACCACTTTTATCATGAAAGTGTCCAC-----CTTTATCCAACT 17209
 QY 7176 CTTCACTGAGAGAAATTTTCAAGGCCAGATTAAGTATATCATGAGGTTCACAGACTG 7235
 DB 17210 CTTACACTGAGAGAAATTTTCTGACAGACCGAGTATATCTTGTATTCACAGCCCTG 17268
 QY 7236 GCTATGATGAGTGGGCTGTTTGTATTTCTTATTTATTTTCCCATGTCACAGAACTTTGTG 7295
 DB 17269 GCTATGATGAGTGGGCTGTTTGTATTTCTTATTTATTTTCCCATGTCAGTAAATTTGTG 17328
 QY 7296 TGCCAGTATGAGAGAGGTGAGGCTTCAAGTCAACAAAGTGGATTTTCAACTGATTTG 7355
 DB 17329 TTGCATATAGAGAGGCCAGGCTTTTATGATCATATATTTTCCCATGTCAGTAAATTTGTG 17388
 QY 7356 AGGAC-CCCACTTTTATGATGATGATCTTATTTCTGTGAGTCTGTGATCTCTCTTT 7414
 DB 17389 AGCACTCTCACTTTTATGAGAGTGTATTTCTGTGAAAGCTGTCTCTCTGTA 17448
 QY 7415 AATGAGACATGTAATTTCCCATGAGAGGTGTGGAGATATGATATCAACACT 7474
 DB 17449 AATGAGAGCAATATTTCTATATATATATATATATATATATATATATATATATATAT 17507
 QY 7475 GGTGATCATCATCTGTTTCTGTTTCAAGGCTCATGAGCTGAGGTTTCTGACATGATTT 7534
 DB 17508 GCTGATCATCATCTTTTGTGTTTGTGTTTCCAGGAGACAGACTAGGTTTCTGTGATGATC 17567
 QY 7535 CAACATCCCACTGTTGGTATGAGAACTGACCAATCAACGAGAGTGAAGAGAC----- 7589
 DB 17568 CAACATCCCACTGTTGGTATGAGAACTGACCAATCAATCAATGAGAGAGAGAGAGAGAGAG 17627
 QY 7590 ----TCTTGTATCAAGAGACCCCTGAGCTTCAAGGAGTGTGATCTGTTTCTCTG 7645
 DB 17628 TTCAATTTGTGAGAGAGATATGATCTTATGTTGTGCTCTCATCTGTGACCTG 17687
 QY 7646 TCGGCTGACAGAGAAAGCGGT--TGCTGTGCTGCTCTGAGCTGAGCATGAGAGAA 7703
 DB 17688 TCGAGCTGAGAGAAATGAGATCATTTCTTGTGCTCTGAGATTTCCGATGACAGAGAA 17747
 QY 7704 CGCTGTCTCATCTTACATCTTCAACCTGTGTGCGGCGCACTTCTTCTTGAAGGCGCA 7763
 DB 17748 CACCTTCTCCCTTACACCTCAACCTGTGCGGCGGCGCACT--TCTTCTGTGCTCCA 17804
 QY 7764 CATTAATTTGCGGCTTACGCTCATCAATATCCGCCATTTCCATTTCCAAATCTCTAG 7823
 DB 17805 GATTTTGAATTTGTAATTTTCTACCATGATCTTCTTCTTCACTTCACTTATTTTAC 17864
 QY 7824 TCTGTGATGACCTTTTCCCTTATTTATAGGCTTAAGCATGAGCGCATGAGAGCGA 7883
 DB 17865 CACTGTGATGACCTTTTCTTATCTTATACGCTGAGCATGAGTGTGAGCTCATGAGAGCA 17924
 QY 7884 GCGCTGCTGTTCATCTGTGTGCGCATCTGTGATCAACATGCGCGCGCGCGCATATCTGTC 7943
 DB 17925 GCACCTGCTGTTCATCTGTGTGCGCATCTGTGATCAACATGCGCGCGCGCATATCTGTC 17984
 QY 7944 ATGCGATGATGATCTGTGTGCGCGCATCTGTGATCAACATGCGCGCGCGCATATCTGTC 8003
 DB 17985 AGGAGTC--GTGTGTCTCTGTGCGCGCATCTGTGATCAACATGCGCGCGCATATCTGTC 18042
 QY 8004 GTTCTGTGATCTTCTGTATTTAGTGTGCTGATTTCTGTTGTGATGAAGCTCATGATTTAT 8063
 DB 18043 GTTCTGTGATCTTCTGTATTTAGTGTGCTGATTTCTGTTGTGATGAAGCTCATGATTTAT 18102
 QY 8064 TACAAATGCGGTGCTGTTTATTTATGTTGTGTTCTGTGTGAGTCAAGCTGTCTGCT 8123

Db 18103 CACAGTCACATGCGTATTTTATATGATGTTCTCTGCGGGTCCAGCCCGCTCTGCT 18162

Qy 8124 GGTGAGATTTCTGTGATCTCCGAAAGATCCCTGACACAGCTGTACGTACCATCT 8193

Db 18163 GGTGAGATCTTGTGATCTCCGAAAGATCCCTGACACAGCTGTACCATCT 18222

Qy 8184 CTTGAGATCTTGTGATCTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8243

Db 18223 GCTGAGATCTTGTGATCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 18282

Qy 8244 TTCCAGATCCACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8303

Db 18283 TTTCGATCCACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18328

Qy 8304 CTTGCTGCTTGTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8363

Db 18329 CTTGCTGCTTGTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18388

Qy 8364 GCGAGCTCAAAATGAGGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8423

Db 18389 GCGAGCTCAAAATGAGGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18448

Qy 8424 TGAGTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8483

Db 18449 TGAGTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18508

Qy 8484 ATTGAGCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8539

Db 18509 ATTGAGCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18568

Qy 8540 TGGCCTGCGACCTTGAACATTAATGATTTTCTTACCTTCTGCTGCTGCTGCTGCT 8599

Db 18569 TGGCCTGCGACCTTGAACATTAATGATTTTCTTACCTTCTGCTGCTGCTGCTGCT 18627

Qy 8600 CAGGCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8618

Db 18628 GTGATTCCTCAATATCTT 18646

RESULT 10
AC107948/c 156839 bp DNA linear PRI 19-JUN-2002
LOCUS Homo sapiens chromosome 11, clone RP11-81D23, complete sequence.
AC107948
AC107948.7 GI:21909529
VERSION
KEYWORDS
HTG.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 156839)
Bliren, B., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone RP11-81D23
Unpublished
2 (bases 1 to 156839)
Bliren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gaidyna, S.,
Garde, S., Goid, S., Goyette, W., Hume, W., Horton, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hume, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Straus, N., Subramanian, A., Talmas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (24-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 156839)
Bliren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepe, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gaidyna, S., Goid, S., Goyette, W., Hume, W., Horton, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hume, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talmas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 19, 2002 this sequence version replaced gr:21307060.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24483
Center clone name: 81_D_23
----- Location/Qualifiers
1. 156839
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

repeat_region	/chromosome="11" /mag="11" /clone="RP11-91D23" /clone_lib="RPC1-11 complement(1..80) /rpt_family="L2" complement(484..779) /rpt_family="A1uSx" complement(794..1094) /rpt_family="A1uSx" complement(1100..1366) /rpt_family="A1uSg" 1435..1692 /rpt_family="A1uDo" complement(1711..1939) /rpt_family="M1R" complement(1245..2344) /rpt_family="A1uSg" 2346..2384 /rpt_family="A1uSg" complement(12791..3083) /rpt_family="A1uSx" 3690..3863 /rpt_family="L2" 5060..5270 /rpt_family="A1uSg" 5900..6027 /rpt_family="PRAM" 6504..6562 /rpt_family="(CCCG)n" 7387..7502 /rpt_family="L2" 7738..7826 /rpt_family="M1R" complement(8164..8273) /rpt_family="L1MC4" 8379..9017 /rpt_family="L3" 10028..10067 /rpt_family="(TG)n" 10386..10494 /rpt_family="L2" 11713..1167 /rpt_family="L1R56" 11768..12065 /rpt_family="A1uSx" 12066..12116 /rpt_family="L1R56" complement(12117..12389) /rpt_family="A1uY" 12390..12706 /rpt_family="L1R56" complement(12901..13000) /rpt_family="L1R49" complement(13001..13348) /rpt_family="A1uDo" complement(13350..13442) /rpt_family="L1R48B" complement(13460..13680) /rpt_family="L1R48" complement(13681..13989) /rpt_family="A1uSg" complement(13990..14172) /rpt_family="L1R48" complement(14305..14325) /rpt_family="L1R23" 14326..14634 /rpt_family="A1uDo" complement(14635..14757) /rpt_family="L1R23" complement(14759..15047) /rpt_family="A1uSx" complement(15090..15395)	Humata
---------------	---	--------

		/rpt family="AluJb"			
	repeat_region	complement(15396, .15689)			
	repeat_region	/rpt family="AluSc"			
	repeat_region	complement(15719, .15812)			
	repeat_region	/rpt family="FRAM C"			
	repeat_region	complement(15814, .16103)			
	repeat_region	/rpt family="LTR23"	/		
	repeat_region	17513, .17811			
	repeat_region	/rpt family="AluSx"			
	repeat_region	complement(18003, .18119)			
	repeat_region	/rpt family="MER81"			
	repeat_region	complement(19303, .19604)			
	repeat_region	/rpt family="AluJb"			
	repeat_region	complement(19814, .19962)			
	repeat_region	/rpt family="FRAM"			
	repeat_region	complement(20403, .20461)			
	repeat_region	/rpt family="L3"			
	repeat_region	complement(20478, .20790)			
	repeat_region	/rpt family="AluJb"			
	repeat_region	21558, .21593			
	repeat_region	/rpt family="(T)n"			
	repeat_region	complement(22105, .22314)			
Query Match	17.8%;	Score 1355.8;	DB 9;	Length 156839;	
Best Local Similarity	91.2%;	Pred. No. 0;			
Matches 1699;	Conservative	0; Mismatches 152;	Indels 12;	Gaps 6;	
QY	6750	ATTCTTCACACAGATGAGGTGGCCCAATAAGATTGCGATGAATAATTTCTT			6809
Db	156837	ATTCTTCACACAGATGAGGTGGCCCAATAAGATTGCGATGAATAATTTCTT			156778
QY	6810	CATCTAATATATATGTGTGTCTTAATGCCCTTACATTGTGCCTTCAATTTTATTTTCAT			6869
Db	156777	CATCCAGTAAAGTGTTGTGTCTTAATGACTTACATTGATGTGCTTCAATCTTATTTTCAT			156718
QY	6870	TTCATCCAAATCTACCATTTGCCCTTATGGCTTCATGATGCAATCTCTTCATTTGAATGAA			6529
Db	156717	TTCATCCAAATCTACCATTTGCCCTTATGGCTTCATGATGCAATCTCTTCATTTGAATGAA			156661
QY	6930	CGTTTATGAAAAGCACATTTGTGCTGTATNGG-AAATAGGACAATAGATATAAATGTAA			6988
Db	156660	TGTATATGAAAAGCACAGTGTGCTGTCTATGAAATAGGACCTGGAGATATATAATGTAA			156601
QY	6989	AATGTGTCCTGTCTGCAATGACTGACACACTGAGTTATTTCTACACCACAGSTCCGC			7048
Db	156600	AATGTGTCCTGTCTGCAATGACTGACACACTGAGTTATTTCTACACCACAGSTCCGC			156541
QY	7049	CATTTTCACAC--ATCTTAGCGAAGATCCCATTTTCTCTGGTTCAATATGATGATCTT			7106
Db	156540	CATTTTCACACTTATCTTAGCAAGAGTCAATTTTCTCTAGTTCAATATGATGATGATCTT			156481
QY	7107	TTTTTCTGTCCAGATGACAGTCTGTGATGAGAGGTGTCACACACCCTTTGTGT			7166
Db	156480	TTTTTCTGTCCAGATGACAGTCTGTGATGAGAGGTGTCACACACCCTTTGTGT			156421
QY	7167	ATCTGAATTCCTCCACCTGAGAGAAATTTTCAGGCCGAGATGAGATTAATCATCGGGTCC			7226
Db	156420	ATCTGAATTCCTCCACCTGAGAGAAATTTTCAGGCCGAGATGAGATTAATCATCGGGTCC			156361
QY	7227	ACAGCACTGGCTAGATGAGTGGGGGTGTTTGAATCTTAATGTTATCCCATGTACAGACA			7286
Db	156360	AAAGCCCTGGCCGAGATGAGTGGGGGTGTTTGAATCTTAATGTTATCCCATGTACAGACA			156301
QY	7287	GAACTGTGTGGCAGTAGAGAAGGTCAAGCTTCAAGATTCACACACAAACCTGGATTTTCAAA			7346
Db	156300	GAACTGTGTGGCAGTAGAGAAGGTCAAGCTTCAAGATTCACACACAAACCTGGATTTTCAAA			156241
QY	7347	CTGAGATTGAGACCCCCACTTTTATATAGAGTCTTATCTCTGGAGTCTGTG-ATCT			7405
Db	156240	CTGAGATTGAGACCCCCACTTTTATATAGAGTCTTATCTCTGGAGTCTGTG-ATCT			156182
QY	7406	CTCTCTTTTAAATGAGACAGTAAATCCACATGACAGGGTGTGGGAGAAATCAAGAT			7465

Db 156181 CTCTTCTTTAAATGAGACAGTAATCCATACGGCAGGGTGTGGGAGATACAGAT 156122
 Qy 7466 CAACACCTGGTGTATCAATCTGGTTTCTGTTTCAGGGTGTATGAGACTGGGGTTTCTGA 7525
 Db 156121 GATACAGCTGGTGTATCATCTGGTTTGTGTTTCCAGGGGACAGAGTACAGGTTTCTGA 156062
 Qy 7526 GCATGATTCACCACTCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCTGAGG 7585
 Db 156061 GCATGATTCACCACTCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCTGAGG 156002
 Qy 7586 AGACTCTTGTATACAGACAGACCTGAGCTTACAGGGGCTGAGGAGATGGTTTCCCTTG 7645
 Db 156001 AGACTCTTGTATACAGACAGACCTGAGCTTACAGGGGCTGAGGAGATGGTTTCCCTTG 155942
 Qy 7646 TCGGCTGACAGAAACGCGGTTGTGCTGTGGCTCTGGGGCTGCGGATGCGAGAAACG 7705
 Db 155941 TCGGCTGACAGAAACGCGGTTGTGCTGTGGCTCTGGGGCTGCGGATGCGAGAAACG 155882
 Qy 7706 CTGCTCTCATCTATCATCTCAACCTGGTCCGGCCGACCTTCTTCTTACGGCCACA 7765
 Db 155881 CTGCTCTCATCTATCATCTCAACCTGGTCCGGCCGACCTTCTTCTTACGGCCACA 155822
 Qy 7766 TTATATGTTGGCCGTTAGGCTTCAATATCCGCAATCCCATGCCAAATTCCTCACTC 7825
 Db 155821 TTATATGTTGGCCGTTAGGCTTCAATATCCGCAATCCCATGCCAAATTCCTCCTT 155762
 Qy 7826 CTGATGACCTTTCCCTACTTTATAGGCTTACAGCTGAGGCGCCATCAGACACGAGC 7885
 Db 155761 CTGATGACCTTTCCCTACTTTTACAGGCTGAGGCGCCATCAGACACGAGC 155702
 Qy 7886 GCTGCTCTCATCTCTGTGGCCCATCTGTGTACATCTGCGCCGCCAGATTCCTGTAT 7945
 Db 155701 GCTGCTCTCATCTCTGTGGCCCATCTGTGTACATCTGCGCCGCCAGATTCCTGTAT 155642
 Qy 7946 CGGTCATGTGTCTGTGCTGCGGCCCTGCTCCCTGCTGCGGAGTATCCTGAGGAGTGT 8005
 Db 155641 CGGTCATGTGTCTGTGCTGCGGCCCTGCTCCCTGCTGCTGTTAGTATGTGAGAGTGT 155582
 Qy 8006 TCTGTGACTTCTCTGTTTATGTGTCTGATCTGTTGTTGTTGTTGTTGTTGTTGTT 8065
 Db 155581 TCTGTGACTTCTCTGTTTATGTGTCTGATCTGTTGTTGTTGTTGTTGTTGTTGTT 155522
 Qy 8066 CAATCGGCTGCTGCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8125
 Db 155521 CAATCGGCTGCTGCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 155462
 Qy 8126 TCAGAGTCTCTGTGATCCCGAAGATGCGGCTGACCAAGCTGATGAGTCAATCTCC 8185
 Db 155461 TCAGAGTCTCTGTGATCCCGAAGATGCGGCTGACCAAGCTGATGAGTCAATCTCC 155402
 Qy 8186 TCAGAGTCTCTGTGATCCCGAAGATGCGGCTGACCAAGCTGATGAGTCAATCTCC 8245
 Db 155401 TCAGAGTCTCTGTGATCCCGAAGATGCGGCTGACCAAGCTGATGAGTCAATCTCC 155342
 Qy 8246 CCAGATCCCACTGATGAGAAAGTCTTATTTGTGATGCACTGATGTTTCCATTTTCC 8305
 Db 155341 CCAGATCCCACTGATGAGAAAGTCTTATTTGTGATGCACTGATGTTTCCATTTTCC 155282
 Qy 8306 TGTCCGCTTTAAGAGAGTCCCAACCCCATCTTACTTCTGTGGGCTCTTTAGGC 8365
 Db 155281 TGTCCGCTTTAAGAGAGTCCCAACCCCATCTTACTTCTGTGGGCTCTTTAGGC 155222
 Qy 8366 AGGCTCAAAATAGCAGAACTGAGCTGTTCTCAGAGGCTCTGACAGACAGCCTG 8425
 Db 155221 AGGCTCAAAATAGCAGAACTGAGCTGTTCTCAGAGGCTCTGACAGACAGCCTG 155162
 Qy 8426 AGGCTGATGAAGGAGGAGGCTGCTCTCAGAGAACTGAGCTGCTGAGAGAGCAGAT 8485
 Db 155161 AGGCTGATGAAGGAGGAGGCTGCTCTCAGAGAACTGAGCTGCTGAGAGAGCAGAT 155102
 Qy 8486 TGAGAGTGAAGAGAACTCTGCGCT----GTGACAGAGACTTTGAGAGCAATGCTG 8541
 Db 155101 TGAGAGTGAAGAGAACTCTGCGCTGTGACAGAGACTTTGAGAGCAATGCTG 155042

Qy 8542 CCTGCCACCTTGACAAATTATGATTTTCTTACCTTCTGCTTCAGAAATGCTCA 8601
 Db 155041 TCTGCCACCTTGACAAATTATGATTTTCTTACCTTCTGCTTCAGAAATGCTCA 154982
 Qy 8602 GGG 8604
 Db 154981 GTG 154979

RESULT 11
 AC079120/c 104547 bp DNA linear PRI 09-JAN-2002
 LOCUS Homo sapiens BAC clone Rpl1-345M24 from 2, complete sequence.
 AC079120
 AC079120.6 GI:15638822
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 104547)
 Sultson,J.B. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE
 PUBMED
 9847074
 2 (bases 1 to 104547)
 Kalicki,J., Cotton,M. and Elliott,G.
 The sequence of Homo sapiens BAC clone Rpl1-345M24
 Unpublished (2001)
 3 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (18-AUG-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (18-SEP-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 18, 2001 this sequence version replaced gi:14488391.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.edu
 Summary Statistics
 Center project name: H_NH0345M24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.

McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPI1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, B., Tareno, M., Caranese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletier de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: PBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-799L6; the clone sequenced to the right is RP11-551D18, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-345M24; actual end is at base position 78839 of RP11-551D18.

Data from AC068035 was used to finish AC079120. Polymorphisms exist between RP11-345M24 and RP11-799L6.

Location/Qualifiers

```
FEATURES
source
    1..104547
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="2"
        /map="2"
        /clone="RP11-345M24"
        /clone_1kb="RP11-11"
        /size="703..733"
repeat_region
    /rpt_family="AT_rich"
    1163..2264
    /rpt_family="L1"
repeat_region
    2265..3187
    /rpt_family="L1"
repeat_region
    3181..3210
    /rpt_family="AT_rich"
    4072..4465
    /rpt_family="L1"
repeat_region
    4466..4781
    /rpt_family="Alu"
    4757..4781
    /rpt_family="Alu"
repeat_region
    4782..5696
    /rpt_family="(A)n"
    9185..9497
    /rpt_family="Alu"
repeat_region
    9467..9497
    /rpt_family="(CAA)n"
    10035..10067
    /rpt_family="AT_rich"
    10824..10999
    /rpt_family="MaLR"
    11052..11571
    /rpt_family="L1"
    11645..11940
    /rpt_family="Alu"
repeat_region
    12037..12260
    /rpt_family="CR1"
    12570..12968
    /rpt_family="L1"
repeat_region
    12969..13444
    /rpt_family="L1"
    13475..13726
    /rpt_family="Alu"
repeat_region
    13755..13944
    /rpt_family="MERS3"
    14336..14671
    /rpt_family="MaLR"
    14672..15008
    /rpt_family="ERV1"
```

```
repeat_region 15109..15278
    /rpt_family="MaLR"
repeat_region 15374..15405
    /rpt_family="AT_rich"
repeat_region 15583..15712
    /rpt_family="MIR"
repeat_region 16034..16118
    /rpt_family="MIR"
repeat_region 16411..16613
    /rpt_family="MER1_type"
repeat_region 16538..16573
    /rpt_family="AT_rich"
repeat_region 16947..17133
    /rpt_family="MERS3"
    17353..17573
    /rpt_family="L1"
repeat_region 19151..19275
    /rpt_family="L2"
repeat_region 19283..19484
    /rpt_family="MIR"
    20165..20524
    /rpt_family="MaLR"
repeat_region 20620..20655
    /rpt_family="CT_rich"
    20804..21233
    /rpt_family="L1"
repeat_region 21244..21395
    /rpt_family="MIR"
    21540..21691
    /rpt_family="MIR"
    22403..22489
    /rpt_family="CR1"
    22667..23099
    /rpt_family="MaLR"
    23338..23422
    /rpt_family="CR1"
    23429..23604
    /rpt_family="MER1_type"
    23620..23666
    /rpt_family="MER1_type"
    24022..24462
    /rpt_family="MaLR"
    24747..24816
    /rpt_family="L1"
    24801..24835
    /rpt_family="AT_rich"
    24877..24955
    /rpt_family="MIR"
    26330..26535
    /rpt_family="ERV1"
    26588..26638
    /rpt_family="AT_rich"
    27280..27455
    /rpt_family="L1"
    28101..28267
    /rpt_family="Alu"
    28792..29078
    /rpt_family="L1"
    29946..30319
    /rpt_family="MaLR"
    30378..30489
    /rpt_family="L1"
    30511..30670
    /rpt_family="L1"
    34615..34848
    /rpt_family="L1"
```

Query Match 17.7%; Score 1526.8; DB 9; Length 104547;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2047; Conservative 0; Mismatches 367; Indels 140; Gaps 16;
QY 1 TGTATGACCAATGCTACTTTAAT-ACCAAAACGAGAAAGATATA--CAAAAAGAA 57
|||||

Db	85598	CTATTGAAGCAGATATCTCCCTAATAATCAAAATCCAGGAGAGACATAATTAATAAGAA	85539
Qy	58	AACATATGACCAATGACCACTGATGATATATACATGCAAGAAATCCCAACAAATACTAGCT	117
Db	85538	AACATACAGACCAATATATCCCTGATGAAATATGATGACAAAGTCTTCGGCAAAATACAGCT	85479
Qy	118	AACCCATTCACAGCAATATCCAGAAATATATCCACATTGTCTCAATGGGTTTCATACCA	177
Db	85478	AACCAAAATCCAAACGATATATCAAAAAGATATATCCACATGATCAATGTGAATTTCTATATCA	85419
Qy	178	GGGGTGCA-GGATAGTTTAAACATACATACAGATCAATATATGATATCATCACTAAACAGA	236
Db	85418	GGGATGAGGGATAGTATTTATTTATCTCAAGTCAATATATGATACCAATATTAACAGA	85359
Qy	237	ATTAAAACAAAAAATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTTGAACAAATC	296
Db	85358	ATTAAAACAAAAAATCAATGATCATCTCAATAGATGAGAAAAACATTGACAAACAT	85299
Qy	297	TAACTTTCTTTATGATTTAAACCTTGACAAATGACATAGAAAGACATTCCTTAAT	356
Db	85298	CAGATGGCTTTATGATTTAAACCCCTCAGCAAAATTTGCAAGAAAGGACATAGCTTAAG	85239
Qy	357	GTAATTAAGCCATATATGACGAGCCACAGCAAACTTATATGAAATGGGAAAAAGTTG	416
Db	85238	GTAATTAAGCCATCTATGACAAACCCACAGCCACATTAATCTTATGAGAAAAAGTTG	85179
Qy	417	AAAACTATGTCCTGAGAACTGAAACAGACAGATG-CTACTTCAACACTTCTAATTC	475
Db	85178	AAAGCATTTCCCTGAGAAACAGGAAACAGACAGATGTCATCTTCAACAATTCTAATTC	85119
Qy	476	AACTATGATGAGGAATTTTATGCCAGACGATCAGACAGAGAAAGAAATCAAGGCAAC	535
Db	85118	AACCTAGTACTGGAAAGTCTTAGCCAGTGCATCAGACAGAGAAAGAAATCAAGGCAATC	85059
Qy	536	CAATCAATAAAGGAGAGTCAAACTGTCCTCTGTCATCTGATCATATGATTTATATCTTA	595
Db	85058	CAATTTGTAAGAGGAGTCAAACTGTCCTGTTTGCATAGACATTAATCTATATACATA	84999
Qy	596	GAAATCCCTTAAGACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTT	655
Db	84998	GAAATCCCTTAAGACTCATCCAGAAAGCTCTAGAACTGTAATGATGATTCAGTAAAGTT	84939
Qy	656	TCAGGATACAACTAAATGTACACAAATCAGTACCTGCTATACCAACAGTACCA	715
Db	84938	TCAGGATACAAATTAATGTACACAAATCAGTACCTGCTATATACCAACAGTACCTA	84879
Qy	716	GCTGAGAAATCAATCAAGAACTCAAAACCTTTTACATAGCTGTAA-----AAATAC	769
Db	84878	GCTGAGAAATCAATCAAGAACTCAAACTCTTTTACATAGCTTACAAAAATCAATAAATAC	84819
Qy	770	TTAAGATATTTCTTACCCAGAGAGGTGAGAGACCTCTACAAAGAAATCTACAAACACAG	829
Db	84818	TTAGGAATATATCTTACCAAGAGAGGTGAGAGACCTCTACAGAAAACTACAAACACATG	84759
Qy	830	CT-----GACATCATAGTACACAAACAACTGAAAAACATCCATGCTCATAGATGGGT	885
Db	84758	CTGTTAAGAAATCATAGTACACAAACAAATGAAAAACATCTCTATGCTCATGGCTGGGT	84699
Qy	886	AGATCAATATTTTGAAAAATGACCATATTTGCCAAAGCAATCTTACAGTTCAATCAAT	945
Db	84698	AGACTCAATATTTTGAAAAATGACCATATCTGACAAAGCAATCTTACAAATTCATCAAT	84639
Qy	946	CCGACCAAAATATATCATCATCTTTCACAGACTGTAGAAAAAAACAAATCTTAAATTCAT	1005
Db	84638	TCATTCAAATATACACATCATCTTTCACAGAACTAG-AAAAAAACAACTTAAATTCAT	84580
Qy	1006	ATGGAACAAACCAAAAAAAAAAAAAAAAAAAACCCGATGACCAAGCAAGACTTAGCAAA	1065
Db	84579	GTGGAAC-----CAAAAAAAAAAGGCGCTCATAGCCAAAGAAAGACTCAGCAAA	84533
Qy	1066	AAAGAACAAAATCTGAGGATCATCATTAACCATTTTCAAACTATATCTACAGGCTTAAATC	1125
Db	84532	AAAGAACAAAATCTGAGGATCATCATTAACCATTTTCTTAAATATATCTATTAAGCCATAGTC	84473

QY	1126	CCAAAACATCATGCGCACTGACATTAATAAACTGAGGACATACCAATGGAAAAAAGAGAG	1185
Db	84472	ACCAAAACAGCATGTACTGTGTTTTAAAAATAGAAATTAAGATCAATGAGAACAAATAGAT	84411
QY	1186	AATCCAGAAATTAAGCCMAATTAATTATGCGCACTGATTTTGTGACAAAGCAAAACAAAC	1245
Db	84412	AATCCAGAAATTAATCCAAATTACTTAATAGTCAACTGATCTTTGACAAAGTGAACAAAC	84355
QY	1246	ATTAAGTGGGAAAAAGACATTCTTAGTTAAACAAATGTGTCTGAGATTATTGGCAAGCCACA	1305
Db	84352	ATTAAGTGGGAAAAAGATATCTTAATTCAACAGGTGTGGGATTAATTGACAAAGCCACA	84293
QY	1396	TGTGGAAGAAATGAACACTGATCCCTTGTCTCTGCACTTAATACAAAAATTTGATCAAGATG	1365
Db	84292	TGTGGAAGAAATTAATCTGTGGT - CCTCATCTCTCACTTATACAAAATCAACTCAAGAGA	84233
QY	1366	GATCAAGAATTAAATCTGAGACCTTAACCATTAATAAAATTTCTAAGAGATTAACATCGAAA	1425
Db	84233	AATCAAGAATTAAATCTTAAGACCTGTAAACATTAATAAAATTTCTAAGAGATTAACATCGAAA	84177
QY	1426	AATGCTTCTAGACATTCACCTTAGGCAAAAGCTTACGGCCCAAGAACCCCAAAATTAATGC	1485
Db	84173	AATGCTTCTAGACATTCCTTAGGCAAAAGCTTACGACCAAGAACCCCAAAACCAATGC	84111
QY	1486	AACAAAAACAAAAATTAATAGATAGACTTAATTAACCTAAAAAGCTTTTGGCCAGACAAA	1545
Db	84113	AACAAAAACAAAAATTAATAGATAGATTCATTAACCTAAAAAGCTTTCGACACAGAAA	84055
QY	1546	AACATCATTTAGCAGAGCAACACACACCCACCGATGAGAGAAAATCTTCAACAACTA	1605
Db	84053	AGAAATATATCGCAGAGATTAAACACACACCCATATAGATGGGA - AAAATCTTTGCCAATCTA	83993
QY	1606	AGCATCTGACTAAGAGACTAATATCCGGAATCCCAAGAAACTCAACAAATCGACAAGAA	1665
Db	83994	TATTTCTGACAAAGGATTTAATATTGAGAAATCTACAGAAATTTCAACAAATCAACAAGAA	83933
QY	1666	GAAGGCAACCAATCCATGAAAGATGGGCTAAGACATGATATAGCAATTTCTCAAAAGA	1725
Db	83934	AAAAGCAACCAATCTCATCAAAAAGTATGTTAAGATATGAAACAGACAGTTCTCCAAAG	83877
QY	1726	AGATATTCAAATGGCCCAACAAACA - GGAAAAAATGCTTAAATCAATCACTAATGATTAGGAAA	1784
Db	83874	ATATATTCAAATGGCCCAACCAACATATATAAAATATGCTCAACATCACTAATATATCCGGA	83811
QY	1785	ATGTAAATCAACACTGTATATGCGATATCCACTTACTCTGCAAGAAATGTCATATTTTAA	1844
Db	83814	ATGCAATTCAAAACCTCAATTCATATACCTTAACTCTGCAAGAAATGTCATATTTTAA	83755
QY	1845	AAATCTAAAAATTAATAGATGTTGGTGGTCTGTGTGTGATTAAGAAACCTTTTCACTGC	1904
Db	83754	AAATCTAAAAATTAATAGATGTTGGCACTGTGTGTGTGAAAGGACATTTTTCACCTGT	83699
QY	1905	TGTTGGGAAATGTAAACTTTCGCAACAACATAATGGAAGAAAGTGTGAAATTTTCTTAAGAA	1964
Db	83694	TGTTGGGAAATGTAAACTTATGTCACCACTATGGAAGAAAGTGTGTGAGGTTTCTTAAGAA	83633
QY	1965	CTAAAGTAGATGACCATTTGATTCAGCAATCCAT - - - - -	2001
Db	83634	CTAAAGTAGATGATACCATTTGATTCAGCAATCCCACTCCTTGATATTTTACCAAGAGA	83577
QY	2002	-----TAAATAT	2008
Db	83574	AAAGTAGTGTATGCAAAAAAAGCACTGCGACATGCAATGATAT	83515
QY	2009	GTATTAATATATATATTTATATAC - - - - -ATGAAT	2040
Db	83514	ATTAATATGTATATATATATATGATCATATATATCAATTAATATATATATTAATGGAAT	83455
QY	2041	ACAATCAGCCCTTAAAAAAGATTAATAATGATGACTTTCACAGCAATCTAGATGGAATTTGG	2100
Db	83454	ACTACTTAGCCATTTAAAAAAGAAATGAATATATGCAATTTTCAGCAATCTGATGGAATTTGG	83395

QY 2101 AGACCTTATTCTAAGTGGGTAAGTACAGAAATGAAACCAACATCATATGTTCTCAC 2160
 DB 83394 GGACCATATTTTAAAGTGAATTAACAGAGTGGAAAACCAACATCATATGTTCTCAC 83335
 QY 2161 TTACAAGTGGGGCTTAAGTGTGAGACAGAAAGCATAGATATATATGAACTTGG 2220
 DB 83334 TCATATAGTGGAGCTAAGCTATAGAGACAAAGCAT--AAGATATCTATGAGCTTTGG 83277
 QY 2221 GGACTTGAAGGGAAGATGGAAGAGGGAGGAGATTAAGATACATAGGTTACAGT 2280
 DB 83276 GGA-TTGAGGGGAATGAGTGGAGGGGGTGAAGGATTAAGATACATAGGTTACAGT 83218
 QY 2281 GTACACTGCTCAGTGTAGTGGTGCACCAAAATCTCAAAATTAACAATAAGAACTTATC 2340
 DB 83217 GTACACTGCTCGGTGTGTAGTGGTGCACCAAAATCTCAAAATTAACAATAAGAACTTATC 83158
 QY 2341 CATGGAAGCAACACCACTG-TTCCCAAAATCCCAATGAATTAATAATATATATAT 2399
 DB 83157 CATGTAACCAATATACATCTGCTTTCCCTAAACCTATTAATAATTAATAATAATA 83098
 QY 2400 AATGATTTAATTTCACAGATTTAAAGTTCA 2433
 DB 83097 CAATGGCCAGAAACATATGAAAAAGCTCA 83064

RESULT 12
 AC074016/c 179216 bp DNA linear HTG 09-AUG-2001
 LOCUS Homo sapiens chromosome 2 clone RP11-155P18, WORKING DRAFT
 DEFINITION
 AC074016
 AC074016.5 GI:15144340
 HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 179216)
 Waterston, R.H.
 The sequence of Homo sapiens clone
 Unpublished
 2 (bases 1 to 179216)
 Waterston, R.H.
 Direct Submission
 Submitted (10-JUN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Aug 9, 2001 this sequence version replaced gi:9958173.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H.NH0155P18

----- Summary Statistics -----
 Sequencing vector: M13, 488
 Sequencing vector: plasmid, 448
 Chemistry: Dye-terminator Big Dye, 448 of reads
 Assembly program: Phrap, version 0.990319
 Consensus quality: 178616 bases at least Q40
 Consensus quality: 178729 bases at least Q30
 Consensus quality: 178811 bases at least Q20
 Insert size: 174000; agarose-fp
 Insert size: 181802; sum-of-ctnigs
 Quality coverage: 15.53 in Q20 bases; agarose-fp
 Quality coverage: 14.71 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 82838: contig of 82838 bp in length
 * 82839 179216: gap of unknown length
 * 82939 179216: contig of 96278 bp in length.
 Location/Qualifiers

FEATURES
 source
 1..179216
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /clone="RP11-155P18"
 misc_feature
 1..82838
 /note="assembly_name:Contig10
 clone_end:17
 vector_side:left"
 misc_feature
 82939..179216
 /note="assembly_name:Contig11
 clone_end:82838
 vector_side:left"
 ORIGIN

Query Match 17.6%; Score 1521.2; DB 2; Length 179216;
 Best Local Similarity 81.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 383; Indels 82; Gaps 15;

QY 1 TGTATGAGCCAAATGCTACTTTAATACCAAAACGAAAGAGATTA--CAAAAAGAAA 58
 DB 50441 TGTATGAGCCAAATGCTACTTTAATACCAAAACGAAAGAGATTA--CAAAAAGAAA 50382
 QY 59 ACTATGACCACTACCACTGATGATATATATATATATATATATATATATATATATATAT 118
 DB 50381 ACTATGACCACTACCACTGATGATATATATATATATATATATATATATATATATATAT 50322
 QY 119 ACCAATCCAAAGCATATCAAGAT 178
 DB 50321 ACTGATCCAAAGCATATCAAGATATATATATATATATATATATATATATATATATATAT 50262
 QY 179 GGGTGCAGATAGGTTAATACATACAGAGTCAATTAATGTATATATATATATATATATATAT 238
 DB 50261 GGGTGCAGATAGGTTAATACATACAGAGTCAATTAATGTATATATATATATATATATATAT 50206
 QY 239 TAAAAACAAATATCAT 298
 DB 50205 TAAAAACAAATATCAT 50146
 QY 299 ACATTTCTTTATGATTAATACCTTCAGCAAAATGCAATATGAAAGACATATCTTAATGT 358
 DB 50145 GCA-TGCTTTATGATTAATACCTTCAGCAAAATGCAATATGAAAGACATATCTTAATGT 50087
 QY 359 AATTAAGCCAT 418
 DB 50086 AATTAAGCCAT 50027
 QY 419 AATATGTCCTGAGAACTGGAACAAAGAGATG-CTACTTACCACTCTTAATCA 477
 DB 50026 AATATGTCCTGAGAACTGGAACAAAGAGATG-CTACTTACCACTCTCTTAATCA 49967
 QY 478 CATATGATGGAAGTTTATGCAAGCATATGCAAGAGAAAGAAATCAAGGCAACCA 537
 DB 49966 CATATGATGGAAGTTTATGCAAGCATATGCAAGAGAAAGAAATCAAGGCAACCA 49907
 QY 538 AATCAATTAAGAGAAATGAAATGCTGCTGTTCACTGATATATATATATATATATATATATAT 597
 DB 49906 AATCAATTAAGAGAAATGAAATGCTGCTGTTCACTGATATATATATATATATATATATATAT 49847
 QY 598 AATCAATTAAGAGAAATGAAATGCTGCTGTTCACTGATATATATATATATATATATATATAT 657
 DB 49846 AATCAATTAAGAGAAATGAAATGCTGCTGTTCACTGATATATATATATATATATATATATAT 49787

658 AGGATCAAACTAAATGTACCAATGAGTACCTGTATACCAACATGACCAAGC 717
 49786 CAGATCAAAATTAATGTACCAATGAGTACCTGTATACCAACATGAGC 49727
 718 TGGAATCAATCAAGAACTCAACCTTTTACATAGCTGTAAATAATCTA--- 773
 49726 TGAAATCAATCAAGAACTCAACCTTTTACATAGCTGTAAATAATCTA--- 49667
 774 ---GAATATTTCTTACCAAGAGGTGAAGACCTCTACAGAAATACTACAAACAGAC 830
 49666 TTGACATATACCTTACCAAGAGGTGTAGACCTGTACAGAAATACTATTAACCTGC 49607
 831 T---GACATCATATGACCAACCAAGTGGAAACATCCCATGCTCATAGATGGGTAG 887
 49606 TAAAGAAATCATATGACCAACCAAGTGGAAACATCCCATGCTCATAGATGGGTAG 49547
 888 AATCAATATTTGTGAATAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATCC 947
 49546 AATCAATATTTGTGAATAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATCC 49487
 948 CACCAAAATATCATCATCTTCTTACAGAACTGAAATAATCTAATTTCAATAT 1007
 49486 CATCAAAATATCACCATCTGTTTTACAGAACTGAAATAATCTTAAATCATATG 49427
 1008 GGAACACCAACCAAAAAAAAAAAAAAAAACCCCCATGACCAAGCAAGCTTACCAAA- 1066
 49426 GAACCAAAAAAAAAAGAACCAATGACCAAGCAAGCTTACCAAAACAAACCAAAAC 49367
 1067 -----AGAACAAATCTGAGGCAATCATATTACCA 1096
 49366 AAGCAAAACCAAAATCAAAAAAGCAAAACCTTAACTGAGGCAATCATATTACCTG 49307
 1097 TCTTCAACTATATCTACAGGCTTAAATCACCAAAACATCATGCACTGACATAAATCTA 1156
 49306 ATTCAAACTATATATATATGAGCCATGACCAAAACATCATGCTGATTAATATA 49247
 1157 GGCAATATGACCAATGGAAGAGAGATCCAGAAATTAACCAATTAATTAATGCC 1216
 49246 GGCAATATGACCAATGGAAGAGAGATCCAGAAATTAACCAATTAATTAATGCC 49187
 1217 AACTGATTTTGGACAAAGCAAAACAAATTAAGTGGGGAAGAAAGCACTTACTAGTACA 1276
 49186 AACTGATTTTGGACAAAGCAAAACAAATTAAGTGGGGAAGAAAGCACTTACTAGTACA 49127
 1277 AATGCTGCTGAGTATTTGGCAAGCAATGCTGGAAGATGAATGAACTGCTGCTC 1336
 49126 AATGCTGCTGAGTATTTGGCTAGCCCATGTAGAGATGAATGAACTGCTGCTC 49068
 1337 TCACCTTAATACAAATTTGATACAAAGTGAATCAAGACTTAATCTGAGACTTAAACC 1396
 49067 TCACCTTAATACAAATTTGATACAAAGTGAATCAAGACTTAAATCTGAGACTTAAACC 49008
 1397 AATAAAATTTCTAGAGATTAATCATGAGAAATGCTTCTAGACATTCCTTGGGCAAGAC 1456
 49007 AATAAAATTTCTAGAGATTAATCATGAGAAATGCTTCTAGACATTCCTTGGGCAAGAT 48948
 1457 TTCAATGCAAGAACCCAAAGTAAATGCAACAAACCAAAATTAATGATAGACTTAA 1516
 48947 TTCAATGCAAGAACCCAAAGTAAATGATTAATAAAACAAAGATTAATGCTGGGACTTA 48888
 1517 ATTAAATCTAAAGAGCTTTTGGGCAAGCAAAACATCATTTAGCAAGCAACAGCAACC 1576
 48887 ATTAAATCTAAAGAGCTTTTGGGCAAGCAAAACATCATTTAGCAAGCAACAGCAACC 48828
 1577 ACCGATGAGAGAAATCTTCAACAACTTACATCTGCTAAGGACTTAATTCGGGATC 1636
 48827 ACGAGATGAGAGAAATCTTCAACAACTTACATCTGCTAAGGACTTAATTCGGGATC 48768
 1637 CACAAGAACTCAACCAATCAGCAAGAGAAAGCAACCAATCCCATGAAAGAGTGGCT 1696
 48767 TACAACCAAACTTAACCAATTAAGCAAGAAAGAA---CAATTAACCATCAAAAGTGGCT 48710
 1697 AAGGACATGATTAACCAATTTCTCAAAAGAGATATACAAATGCGCAACAAAC--AGGAA 1754

48709 AAGGACATGATTAACCAATTTCTCAAAAGAGATTAATTAATGCGCAACCAATATATAA 48650
 1755 AATGCTTAATCATCATATGATTAAGGAATGAAATTAACACTGTATGCGATACAC 1814
 48649 AATGCTCAACATCATATGATTAAGGAATGAAATTAACCAATGATGATACAC 48590
 1815 CTTACTCTGCAAGATGCTCATTAATTAATAATCTAATAATTAATGATGTTGGGCTC 1874
 48589 CTTACTCTCAAGAAATGCGCATATCAAAAAATTAATAATAATGATGTTGGCATGA 48530
 1875 TGTGATTAAGGAACATTTTACATGCTGCGGGAATGTAATCACTGCGCAACCTA 1934
 48529 TGTGATTAAGGAACATTTTACATGCTGCGGGAATGTAATCACTGCGCAACCTA 48470
 1935 TGGAAACAGTGTGAATTTCTTAAGAACTAATAAGTATGACCATTTGATTCAGCA 1994
 48469 TGGAAACAGTGTGAATTTCTTAAGAACTAATAAGTATGACCATTTGATTCAGCA 48410
 1995 ATCCATTAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2054
 48409 ATTC-----ATATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 48373
 2055 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2114
 48372 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 48316
 2115 AGTGGGTAATCACTGAGATGGAACCAACCAATCATATGTTCTCACTTACAGTGGGGC 2174
 48315 AGTGAATTAATCACTGAGATGGAACCAACCAATCATATGTTCTCACTTACAGTGGGGC 48256
 2175 TAACTGTGAGACCAAGCAT-AGATGATTAATTAATTAATTAATTAATTAATTAATTA 2233
 48255 TAACTGTGAGATGAGTGAAGCAATGAGATTAATTAATTAATTAATTAATTAATTAATTA 48196
 2234 AGGA-TGGAAGAGAGGAGGAGTAAAGACTACAAATGAGTACAGTGTACATGCTCA 2292
 48195 AAGAGTGAAGAGGAGGAGTAAAGACTACAAATGAGTACAGTGTACATGCTCG 48136
 2293 GGTGATGGTGCACCAAAATCTCAGAAATTAACCACTAAGAACTTACATGGAAGCAAA 2352
 48135 GGTGATGGTGCACCAAAATCTCAGAAATTAACCACTAAGAACTTACATGGAAGCAAA 48076
 2353 CACCACTGTTCCCAAAATCCCAATGAATTAATAATTAATTAATTAATTAATTAATTA 2400
 48075 CACCTGTTCCCAAAATCCCAATGAATTAATAATTAATTAATTAATTAATTAATTA 48028

RESULT 13
 AC007256/c 181150 bp DNA linear PRI 07-NOV-2001
 LOCUS Homo sapiens BAC clone Rpl1-575C6 from 2, complete sequence.
 DEFINITION AC007256
 ACCESSION AC007256
 VERSION AC007256.5 GI:14327821
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 181150)
 Sultston,J.E. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 2 (bases 1 to 181150)
 REFERENCES
 Du,F., Laplant,Y., Doeber,A. and Moore,B.
 The sequence of Homo sapiens BAC clone Rpl1-575C6
 Unpublished
 JOURNAL 3 (bases 1 to 181150)
 REFERENCE
 Waterston,R.H.
 Direct Submission
 TITLE Submitted (07-APR-1999) Genome Sequencing Center, Washington
 JOURNAL

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (19-APR-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (20-APR-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
6 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (07-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
7 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (03-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
8 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 7, 2001 this sequence version replaced gi:13677176.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0555C06

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frangon, R., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-156B7; the clone sequenced to the right is RP11-53618, 200 bp overlap. Actual start of this

clone is at base position 1 of RP11-575C6; actual end is at base position 180956 of RP11-575C6.

The sequence from base position 134614 to 134714 is derived from PCR product of RP11-575C6 BAC DNA.

The clone RP11-575C6 contains a transposon in the vector.

FEATURES

source	Location/Qualifiers
1.181150	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-575C6"
	/clone_1fb="RPc1-11"
repeat_region	1..354
	/rpt_family="L1"
misc_feature	91..95
	/note="similar to Homo sapiens EST BG495251 (NID:913456766)"
repeat_region	124..153
	/rpt_family="(TG)n"
repeat_region	192..228
	/rpt_family="(T)n"
repeat_region	330..1055
	/rpt_family="L1"
repeat_region	553..620
	/rpt_family="T-rich"
repeat_region	1056..1363
	/rpt_family="Alu"
repeat_region	1210..1243
	/rpt_family="(T)n"
misc_feature	1228..1246
	/note="match to EST AW444732 (NID:96986494)"
repeat_region	1364..2264
	/rpt_family="L1"
repeat_region	3338..3550
	/rpt_family="MIR"
repeat_region	3739..3866
	/rpt_family="MIR"
misc_feature	3971..4133
	/note="similar to Homo sapiens EST AU117593 (NID:910932573)"
misc_feature	4041..4495
	/note="similar to Homo sapiens EST A1499220 (NID:94391202) to08f03.x1"
misc_feature	4053..4639
	/note="similar to Mus musculus EST BF163999 (NID:91044285)"
misc_feature	4092..4830
	/note="similar to Homo sapiens EST AL528010 (NID:912791503)"
misc_feature	4128..4340
	/note="match to EST AV760501 (NID:910918349)"
misc_feature	4184..4742
	/note="match to EST A1052250 (NID:93308241) oz21c07.x1"
misc_feature	4339..4903
	/note="similar to Homo sapiens EST AU117593 (NID:910932573)"
misc_feature	4485..4491
	/note="similar to Homo sapiens EST BE788113 (NID:910209311)"
misc_feature	4657..5468
	/note="similar to Homo sapiens EST AL563032 (NID:912912044)"
misc_feature	4712..5400
	/note="match to EST AU131936 (NID:910992290)"
misc_feature	4780..5582
	/note="similar to Homo sapiens EST BG164237 (NID:912671000)"
misc_feature	4963..5682
	/note="similar to Homo sapiens EST AU124988

```
misc_feature (NID:g10949704) "
5106..5633
/note="similar to Mus musculus EST A1116325 (NID:g3516649)
ue97c04.y1"
5169..5675
/note="similar to Homo sapiens EST A1144381
(NID:g11005902) "
5194..5761
/note="similar to Homo sapiens EST BF529656
(NID:g11617019) "
5237..6019
/note="similar to Homo sapiens EST A1130508
(NID:g10990862) "
5238..6043
/note="similar to Mus musculus EST BF166846
(NID:g11047198) "
5326..5653
/note="similar to Homo sapiens EST BG390376
(NID:g13283824) "
5361..6051
/note="similar to Homo sapiens EST A1135897
(NID:g10996436) "
5361..5812
/note="similar to EST BG862638 (NID:g14213176) "
5420..6048
/note="similar to Homo sapiens EST BF788113
(NID:g10209311) "
5551..6327
/note="similar to Homo sapiens EST A1130771
(NID:g10991125) "
5615..5984
misc_feature
Query Match 17.6%; Score 1521.2; DB 9; Length 181150;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 383; Indels 82; Gaps 15;
QY 1 TGTATGAAGCCATGTCACTTAAATACCAAAACGAGAAAGATATA--CAAAAAGAA 58
Db 100770 TGTATGAAGCCAGATACACCTTAATACCTMAAAAGAGAAAGATACCAAAAAAGAAA 100711
QY 59 ACTATAGACAGTCACTGATGATATACATGAGAAATCCCAACAAATATCTAGCTA 118
Db 100710 ACTACAGACAAATATCTTGAATGACATGATGCAAAAATCTTTAAACAAATCTGGCTA 100651
QY 119 ACCCAATCCACAGCATATCAAGAATATCAACATTTCAAGTGGGTTTCAATACAG 178
Db 100650 ACTGAATCCACAGCATATCAAGAATATCAACATTTCAAGTGGGTTTCAATACAG 100591
QY 179 GGGTGCAGATAGTTAAATCAATCAAGTCAATTAATGTGATACATCAATAAACAGAT 238
Db 100590 GGAATATAG---GCTTAAATCAATCAAGTCAATTAATGTGATACATCAATAAAGAT 100535
QY 239 TAAAAACAAAATACATGATCATCTCAATGATGCTGAAAAGCATTTGACAAATCTA 298
Db 100534 TAAAAATAAATATCAATGATCATCTCAATGATGCGAAAAGCATTTGACAAATCTCA 100475
QY 299 ACATTTCTTATGATTAATACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGT 358
Db 100474 GCA-TGCTTATGATTAATATCTCAAGCAAAATAGATTAACAAGGCGATCTCAATGT 100416
QY 359 AATTAAGCCATATATGACGACCCACAGCAACATTAATCTGAATGGGAAAAGTTGAA 418
Db 100415 AATTAAGCCATATATGACCAACCCACAGCCAAATCATCTGAATGGGAAAAGTTGGA 100356
QY 419 AACATTTCCCTGGAATGGAACAAGAGATG--CTACTTCAACACTCTTAATCAA 477
Db 100355 AGCATTTCCCACTGAGATGGAACATGACAGATGACGATCTCAACACTCTCTTCAA 100296
QY 478 CATATGATGGAAGTTTATGACGACGATCAAGAGAAAGAAATCAAGGCGACCCA 537
Db 100295 CATATGATGGAAGTTCTAGCAGAGCAATCAAGACAGAGAGAAATTAAGGCGACCCA 100236
QY 538 AATCAATTAAGAGAGATCAAACTGTCTCTTCACTGATGATATGATTGATTAACCTAGA 597
```

```
Db 100235 AATCAGTAAGAGAGAGATCAAACTATTTGCTGTGATGATATTAATGTATACCTACA 100176
QY 598 AAACCTAAAGACATCTCCAGAAAGCTCTTGAACATATCATTAATTCAGTAAGTTTC 657
Db 100175 AAACCTAAAGATTCCTCCAAAAGCTCTTGAACATATTAAGAAATTCAGTAAGTTTC 100116
QY 658 AGGATCAAACTAAATGTACCAAAATGATGACCTGCTATACCAACAGTGACCAAGC 717
Db 100115 CAGATCAAAATTAATGTGACCAAAATGATGCTCTGCTATACCTACAGCGCATAGG 100056
QY 718 TGAATATCAAAATCAAGAACTCAACATTTTACAAATGCTGTAAAAAATCTTAA---- 773
Db 100055 TGAATATCAAAATCAAGAACTCAACCTTTTACAAATGCTGTCAAAAGAAAAATTAATGT 99996
QY 774 ---GATATTTCTTACCAAGAGGTGAGAGACCTCTCAAGAGAAATCAACAAACAGAGC 830
Db 99995 TTGACCTATACCTTACCAAGAGGTGTAAGACCTGTCAAGAGAAATCTATTAACACTGC 99936
QY 831 T--GACATCATGATGACACAAACAAAGTGAAACACATCCCATGCTCATGATGGGTAG 887
Db 99935 TAAAGAAATCATGATGACACAAACAAATGAAACACATCCCATGCTCATGATGGGTAG 99876
QY 888 AATCAATATTTGTGAATATGACCAATTTGCCAAAAGCAATCTAAGAGTTCAATGAATTC 947
Db 99875 AATCAATATTTGTGAATATGACCAATCTGCAAAAGCAATCTAACAATTTAATGCAATTC 99816
QY 948 CACCAAAATATCATCATCTTCTTCAAGAACTGAAATTAATCTAATTAATTCATAT 1007
Db 99815 CATCAAAATATCAACCATCTGTTTTCACAGAACTGAAATTAATCTAATTAATTCATATG 99756
QY 1008 GGAACAAACAAACCAAAAAAAGAAAAAACCCTGATAGCCAAAGCAAGACTTACCAAA- 1066
Db 99755 GAAACCAAAAAAGAACCAACATGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 99696
QY 1067 -----AGAACAAATCTGAGGATCACTTACCA 1096
Db 99695 AAGCAAAACAAAATCAAAAAAGCAAAAAACCTTAACAAATCTGAGGATCACTTACCTG 99636
QY 1097 TCTTCAACTATATCTACAGGCTTATATCAACAAACATCATGAGCATGATTAATTA 1156
Db 99635 ATTTCAAATCTATATTAATGCGCATGATCAACAAACATCATGATTAATTAATTA 99576
QY 1157 GGCACATAGACCAATGGAAGAAAGAGATTCAGAAATTAAGCCAAATTAATTAAGCC 1216
Db 99575 GGCACATAGACCAATGGAAGAAAGAGATTCAGAAATTAAGCCAAATTAATTAAGCC 99516
QY 1217 AACTGATTTTGAACAAAGCAAAACAAACATTAAGTGGGAAAAAGCAATCTTACTTAACA 1216
Db 99515 AACTGATCTTTGACTTAAGCAAAACCAATTAAGTGGGAAAAAGCAATCTTACTTAACA 99456
QY 1277 AATGCTGCTAGATTAATTTGCAAGCCACATGTGGAATGAATGAACTGGATCCTTGTCTC 1336
Db 99455 AATGCTGCTGGAATTAATTTGCTAGCCACATGTAGAGAAATGAATGAT--CTCATCTC 99397
QY 1337 TCACTTAATCAAAAATTTGATACAGATGATCAAGACTTAATCTTGAAGCTTAAACC 1396
Db 99396 TCACTTAATCAAAAATTTGATACAGATGATCAAGACTTAATCTTGAAGCTTAAACC 99337
QY 1397 AATTAATTTCTAAGAGTAACTATGAAAAAATGCTTTAAGACATTCATTTAGGCAAGAC 1456
Db 99336 AATTAATTTCTAAGAGTAACTATGAAAAAATGCTTTAAGACATTCATTTAGGCAAGAT 99277
QY 1457 TTCAATGCAAGAACCCAAAGTAAATGCAACAAACAAATTAATTAATGATGAGACTTA 1516
Db 99276 TTCAATGCAAGAACCCAAAGTAAATGCAACAAACAAATTAATGATGAGACTTA 99217
QY 1517 ATTAACTAAAGCTTTTGGAGCAAAAAACAATCAATTAGCAGAGCAACAGCAACCC 1576
Db 99216 ATTAACTAAAGCTTTTGGAGCAAAAAACAATCAATTAGCAGAGCAACAGCAACCC 99157
QY 1577 ACCGATGAGAGAAATCTTCAACAACTGAATCTGAAGCTTAATATTCGGGAATC 1636
Db 99157 ACCGATGAGAGAAATCTTCAACAACTGAATCTGAAGCTTAATATTCGGGAATC 1636
```

[illegible]

Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dabhone, S. R., David, R., Davila, M. L., Davis, C., Day-Carroll, L., Dedrich, D. A., Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dinh, H. H., Douthett, K. J., Draper, H., Dugan-Rocha, S., Durkin, K. J., Earhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Emertling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Francis, P., Gabel, A., Gao, Y., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guavara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, E., Kovar, C., Kratoch, J., Kurehl, S., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L. J., Li, J., Li, Z., Licharev, O., Lien, C., Liu, J., Liu, W., Louised, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapus, P., Marcondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., MacIntyre, I., McLeod, M. P., Medora, M., Mei, G., Metzger, S., Metker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokemko, S., Ogih, M., Okumura, G., Oragune, N., Oyedro, E., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojoudkan, I., Rolle, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoshetel, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, K., Suton, A. A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Taney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. P., Zhou, J., Zorrilla, S., Kuchelapatti, R., Weinstein, G., and Gibbs, R.

Unpublished

Direct Submission

2 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (08-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

INFORMATION: <http://www.hgsc.bcm.tmc.edu> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:
STRs are identified using ePCR (Genome Reel 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green

Db	122812	AAACATTCTCTGGAAGATGAAA	CAAGA	CAAGAGTCCCTACTTCCAGATCACTCCTCA	122753
OY	475	CACACTAGTAGTGGAGAGTTT	TAGCCGAG	CATCAACAAGAAAAGAAATCAAGGGCAC	534
Db	122752	CAGCATAGTACTAGAAATCCT	TAAACA	CACCAATCAACAAGAAAAGAAATAAAGGGCAT	122693
OY	535	CCAAATCAATAAAGAGGAAGT	CAAACTG	TCCCTGTTCACTGATGATGATGTATGATACCT	594
Db	122652	CAAAATCAATAAAGAGGAAGT	CAAAATTG	TCTGCTGTTGCTGATGATGATGATCACTTTACCT	122633
OY	595	AGAAAACCTTAAGA	CTCATCCG	AAAGCTCCTAGAACTGATACATAAATTCAGTAAAGT	654
Db	122632	CGAAAACCTTAAGA	CTCCTCCG	AAAGCTCCTAGAACTGATGAAAGAAATTCAGCAAAGT	122573
OY	655	TTCAAGATCAAACTAAATGT	ACAAATCA	AGTACAAATCAAGTACCACTGATCAACCACTGACCA	714
Db	122572	TTCTGGATCAAGATATATGT	ACAAATCA	AGTACCAATCAAGTACCACTGATCAAGTACCA	122513
OY	715	AGCTGGAATCAAAATCAAGA	ACTCAACA	CTTTTACATATGCTGTAA-----AAA	764
Db	122512	AG-TGGGAATCAAAATCAAGA	ACTCAACT	CTTTTAAATATGCTGCAAAAATTAATAATAA	122454
OY	765	AATACTTAAAGATATTTCTT	ATACCAAG	AGGTGGAAGCTCTTCAAGAGAAAATCAACAAA	824
Db	122453	AATACTTGAATATATCTTAC	CAAGAGAG	GTGAAGTCTCTTCAAGGGAAGTCAACAAA	122394
OY	825	CACAGCT---GACATCAT	ATGATGACA	CAACAAGTGGAAAACATCCATGCTCATGGA	880
Db	122393	CACCTCTGAAAAGAAATCAT	ATAGATGACA	CAACAAGTGGAAAACATCCCATCTCATGGA	122334
OY	881	TGGGTAGATCAATATTTGTG	AAAAATG	ACCATATGTCGCAAAAGGAATCTCAATTCATG	940
Db	122333	TGGGTAGATCAATATTTGTG	AAAAATG	ACCATATGTCGCAAAAGGAATCTCAATTCATG	122274
OY	941	CAATTCCTCCACAAAATAT	ATCATCAT	CTTCTTCACAGAACTAGAAAAAAACAATTCCTAAA	1000
Db	122273	CAATCCCATCAAAAATCA	CACTATCT	CTTCACAGGTTAG-AAAAACAATTCCTAAA	122215
OY	1001	TTCTATATGGAACAACA	CAAAAAA	AAAAAACCAGCATGACCAAGCAAGCTTA	1060
Db	122214	TTCACTATG-----AA	ACAAAAG	AGAGCCCAATATGCAAGCAAGCAAGCTTA	122170
OY	1061	GGAAGAAAGAACAAATCT	GAGGATCA	CACTTACCATCTTCAAACTATCTAACAAGCTTA	1120
Db	122169	GGAAGAAAGAACAAATCT	GATGATCA	CACTTACCATCTTCAAACTATCTAACAAGCCCA	122110
OY	1121	TATATCAACAAAACAT	CATGGCACT	GTAAACTAGGCACTATGACCAATGAGAAAAGAA	1180
Db	122109	TAGTCAACCAAAACAGAT	GTGACTGT	GTAAACTAGGCACTATGACCAATGAGAAATGAA	122050
OY	1181	GAGAGATCCAGAAATTA	AGCCAAAT	TAATATAGCCAACTGATTTTGAACAAAGCAACA	1240
Db	122049	TAGTGAACCAAGAAATTA	AGCCAAAT	TAATAGCCGCTGTGATCTTCAACAAACCAACC	121990
OY	1241	AAAACTAAAGTGGGAAAA	AGAACTTT	CTAGTTAAACAATGTGTCTGAGATTAATGGCAAG	1300
Db	121989	AAAACTAAAGTGGGAAAA	AGAACTTT	CTAGTTAAACAATGTGTCTGAGATTAATGGCTAG	121930
OY	1301	CCACATGTGGAAGATGA	AACTGATC	CCCTGTCTCTCACTTAATCAAAAAATTTGATACA	1360
Db	121929	CTACAGATGAGAAATGA	AACTGGAAT	-CTCTAGCTCTCACTCTTGTACAAAAGCAACTCA	121871
OY	1361	AGATGGATCAAAAGACTT	AAATCTG	AGACCTTAACCAATTAATAATTTCTAGAGATAACATC	1420
Db	121870	AGATGGATTAAGGCGCTT	AAATCTGA	AGCCTGAAGCTATATAAATTTCTAGAGATAACAT	121811
OY	1421	AGAAAAATGCTTTAGAC	ATTCACTT	ATGCAAAAGACTTCATGACCAAGAAACCAAAAGTA	1480
Db	121810	GGAAAAACCTTTTAA	CAATTTGG	CTTATAGGCACAAAGAAACCAAAAGCA	121751
OY	1481	AATGCAACAAAAACA	AAATTAAT	TAGATGGACTTAATTAATAAAG-CTTTTGGCC	1539
Db	121750	AATGCAATAAAAACA	AGATTAAT	TAGCTGGACCTTAATTAAGCTTAAGAAAGCCTTTTGGAC	121691

QY	1540	AGCAAAAACATCATTAGAGAGGCAACAGCAACCCACGAGTGGAGAAAATCTTCAC	1599
Db	121590	AGCAAAAGGAAACATACGAGAGGTAAACAGGCATCTCCAAAGTGGAGAAAATCTTCAC	121631
QY	1600	AAACTAAGCATCTGACTAAGAATATATCCGAAATCCAGAAAGACTCAACCAATCAG	1659
Db	121630	AAATTATACATCGACAAAGGACTAATATCCAGATCTTCAATGAACTCAACCAATCAG	121573
QY	1660	CAAGAGAAAGCAACCATTCCTCATGAAAGAGTGGGCTTAAGACATGAAATAGCAATTCTC	1719
Db	121570	TAAAGAAAAAAACCAACCAATCCCATCAAAAAGTGGCTAAGACCTGAATAGACAAATCTC	121511
QY	1720	AAAAAGAGATATCAATCAATGGCCCAACAAACAGGAAAAAATGCTTAACATCACTAATGATTA	1779
Db	121510	-AAAGAGATATCAATCAATGGCCCAACAAACATGAAAAATCTCTCTGATCAAGATTTTCA	121452
QY	1780	GGGAAATGTAAATCAACACTGTATGCGATACCACTTACTCTCGCAAGAAATGCTCATTA	1839
Db	121451	GGGAAATGTCAAAATCAAAACCAACATGCAAGTATCTCTCTCTCGCAAGATGTGACATTA	121392
QY	1840	TTTTAAAAATCTAAAAA-TAATAGATGTTGGTGGGCTGTGGTGTAAAGAAACCTTTA	1898
Db	121391	TCGAAAAATCAAAAAACAGTAGCTGTGGTGTGTGATGGGTGATCAAGGAAACCTTCTA	121332
QY	1899	CACGTGCGTGGGAAATGTAAACTTGGGCAACACATCATGAAAAACAGTGTGAAAAATTTCTT	1958
Db	121331	CACGTGCGTGGGAAATGTAAACTAATGACAGCCACATATGAAAAACAGTGGGAGATTCCTT	121272
QY	1959	AAGAACTAAAGTAGATGACCACTTTGATCCAGCAATCCCATTAATATGTATAAATAT	2018
Db	121271	AAAGAACTAAATTAACATCACATTTGATCCAGCAATCCCATTA-----	121222
QY	2019	ATAATTTTATATCAATGGAATCAACCTGAGCAATAAAAAGAAATTAATAGATGACATTC	2078
Db	121226	-----CTGGGTATCTTACCCAGAAAAAGAAAGATTAACAGCATTTT	121182
QY	2079	ACAGCAATCTAGATGGAATTTGAGACCCTTATTTCTAATGGGGTAACTCAGAAATGAAAA	2138
Db	121181	GCAGCAAACTGGGTGAGACTGGAGACTAATTTCTAATGGAAGTGACTCAGAAATGAAAA	121122
QY	2139	ACCAAACTATATGTTTCTCACTTACAGATGGGGCTAAAGCTGTGAGGACACGAAGCAT	2198
Db	121121	ACCAAACTATATGTTTCTCACTGATGTGGAACTAAGCTATGAGGATCAAAAGGCAT	121062
QY	2199	-AGAAATATATATGAACTCTGGGCACTTGAAGGGAAAGATGGAAGAGAGCGAGGGATA	2257
Db	121061	AAGATATATCAATGCACTTTGGGCACTTGGGGGAAAGATGTGAGAGGGCAAGGGATA	121002
QY	2258	AAAGACTACCAATGGGTGCAAGTGTACACTGTCAAGTGTATGGGTGACCAAAATCTTCAG	2317
Db	121001	AAAGGTACAAATAGGGGTGCAATGTATATCTGTTTGGATGATGGGTGCACTAAATCTTCAC	120944
QY	2318	AAATATACCACTAAAGAACTTATTCATGGAAGCAACACACTGTTCCTCCCAAAATCCCA	2377
Db	120941	AAATCACTCACTAAAGAACTGTCAAGTAAACCAAAATATCACCTGTATCCCAAT-ACCTA	120883
QY	2378	TGAATTAATAATATATATATATTAATGATTTATTTTACAGAAATTTAAAAAGTCACTGT	2437
Db	120882	TGAAAAAATATATTAATAAGAAAGAAACAGCTTAAGCATGTGTTTGAGAAAAAGATATTTC	120822
QY	2438	TCAGAGTTTATTAATTAATGAAATGAAATGAAAGTGT	2475
Db	120822	TGAGAGAGAACTTAATGCAAGAACATGATCTGTGA	120785
RESULT 15			
AC006477			
LOCUS	AC006477	121520 bp	DNA linear PRI 27-JAN-2004
DEFINITION	Homo sapiens PAC clone RP4-802615 from 7, complete sequence.		
ACCESSION	AC006477		
VERSION	AC006477.4	GI:14589645	
KEYWORDS	HTG.		

SOURCE	ORGANISM
REFERENCE	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 121520)
AUTHORS	Hillier, L.M., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H., Wagner-McPherson, C., Layman, D., Mias, J., Jaeger, S., Walker, R., Wyllie, K., Sekhon, M., Becker, M.C., O'Laughlin, M.D., Schaller, M.E., Reuel, G.A., Delhanty, K.D., Miner, T.L., Nash, W.E., Cordes, M., Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S., Isak, A., Vambrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L., Kalcicki, J., Ozersky, P., Biellocki, L., Scott, K., Holmes, A., Harkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C., Dauphin-Kohlberg, S., Kozlowski-Reilly, A., Leonard, S., Rohlfing, T., Rock, S.M., Tin-Wollam, A.M., Abbott, A., Mink, P., Maupin, R., Stowman, C., Letreille, P., Miller, N., Johnson, D., Murray, J., Moesener, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Wallis, J.W., Speth, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohlmann, P.E., Cook, L.L., Hickenboehm, M.T., Eldred, J., Williams, D., Bedell, J.A., Mardis, E.R., Clifton, S.W., Chissee, S.L., Marra, M.A., Raymond, C., Hubbard, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadonoto, S., Bubb, K., Stims, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.J., Furey, T.S., Baertsch, R.A., Brent, M.R., Kehler, E., Flick, P., Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrents, D., Chinwalla, A.T., Gish, W.R., Eddy, S.R., McPherson, J.D., Olson, M.V., Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.
TITLE	The DNA sequence of human chromosome 7
JOURNAL	Nature 424 (6945), 157-164 (2003)
MEDLINE	22737999
PUBMED	12853948
REFERENCE	2 (bases 1 to 121520)
AUTHORS	Bauer, C., Joshi, C. and Le, T.
TITLE	The sequence of Homo sapiens PAC clone RP4-802G15
JOURNAL	Unpublished (2001)
REFERENCE	3 (bases 1 to 121520)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	4 (bases 1 to 121520)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 121520)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	6 (bases 1 to 121520)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (03-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	7 (bases 1 to 121520)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	8 (bases 1 to 121520)
AUTHORS	Wilson, R.
TITLE	Direct Submission
JOURNAL	Submitted (27-JUN-2004) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Jul 3, 2001 this sequence version replaced gi:7243871.

	Center: Washington University Genome Sequencing Center
	Center code: MUGSC
	Web site: http://genome.wustl.edu
	Contact: saplens@wustl.wustl.edu

----- Summary Statistics
Center project name: H_DJ0802G15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only provide sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
This clone was derived from human PAC library RPC1-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Iannou et al., *Nature Genetics* 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>), or from Pieter de Jong.
VECTOR: pcYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is CTA-318M5, 200 bp overlap.
Actual start of this clone is at base position 1 of Rp4-802G315.
Actual end is at base position 14906 of CTA-318M5.

FEATURES	location/qualifiers
source	1. .121520
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="RP4-802G15"
	/clone_id="RP4-802G15"
	10. .995
repeat_region	/rpt_family="L1"
repeat_region	996. .1816
repeat_region	/rpt_family="L1"
repeat_region	1975. .2278
repeat_region	/rpt_family="Alu"
repeat_region	2821. .2977
repeat_region	/rpt_family="L2"
repeat_region	4195. .4238
repeat_region	/rpt_family="L1"
repeat_region	5073. .5257
repeat_region	/rpt_family="MIR"
repeat_region	5870. .6457
repeat_region	/rpt_family="L1"
repeat_region	6463. .7323
repeat_region	/rpt_family="L1"
repeat_region	7539. .7638
repeat_region	/rpt_family="CT-rich"
repeat_region	7687. .7857
repeat_region	/rpt_family="MIR"
repeat_region	8517. .8547
repeat_region	/rpt_family="AT-rich"
repeat_region	9124. .9444

```

repeat_region      /rpt_family="L1"
                    9445..9796 "MaLR"
repeat_region      /rpt_family="L1"
                    9797..10969
repeat_region      /rpt_family="L1"
                    10970..11273
repeat_region      /rpt_family="Alu"
                    11274..11395
repeat_region      /rpt_family="L1"
                    11414..12199
repeat_region      /rpt_family="L1"
                    18423..18683
repeat_region      /rpt_family="L1"
                    18676..18921
repeat_region      /rpt_family="L1"
                    18922..19856
repeat_region      /rpt_family="L1"
                    19983..20104
repeat_region      /rpt_family="L1"
                    20105..20470
repeat_region      /rpt_family="MaLR"
                    20471..20591
repeat_region      /rpt_family="L1"
                    20592..20898
repeat_region      /rpt_family="Alu"
                    20902..20925
repeat_region      /rpt_family="(GGA)n"
                    20932..21065
repeat_region      /rpt_family="L1"
                    21079..22678
repeat_region      /rpt_family="L1"
                    22675..22901
repeat_region      /rpt_family="MaLR"
                    22902..24732
repeat_region      /rpt_family="L1"
                    24733..24843
repeat_region      /rpt_family="Alu"
                    24844..25627
repeat_region      /rpt_family="L1"
                    25628..25734
repeat_region      /rpt_family="MER2_type"
                    25768..25823
repeat_region      /rpt_family="CT-rich"
                    25851..26075
repeat_region      /rpt_family="Alu"
                    26076..26684
repeat_region      /rpt_family="MER2_type"
                    26685..26719
repeat_region      /rpt_family="(CA)n"

Query Match      17.5%; Score 1507.4; DB 9; Length 121520;
Best Local Similarity 79.3%; Pred. No. 0;
Matches 2020; Conservative 0; Mismatches 361; Indels 166; Gaps 13;

```

```

300 CATTTCTTATGATTAAACCTTCAGCAAAATCGACATGAAAGGACATACCTTAATGTA 359
    |||
93135 CATCTTTATATATTAACCTTCAGCAAAATCGGACATGAAAGGACATACCTCATGTA 93194
    |||
360 ATAAAAGCATATATGACGGAACCCACAGCAAACTTATATAGTGGGAAAAAGTTGAAA 419
    |||
93195 ATAAAAGCATCTATGACAAACCCACAGTAACTATATGTAAGGGGAAAAAGTTGAAA 93254
    |||
420 ACATTGTCCTGAGAACTGGAACAAAGACAGAGATG-CTACTTCAACCTTATTCAC 478
    |||
93255 GCATTCCTCTGAGAAATGAAACAAAGACAGAGATGCCCTCTCTCACCCTCTTCAC 93314
    |||
479 ATAGTAGTGAAGCTTTTACCGACAAATCAACAAAGAAAGAAATGAAGGGACCCAA 538
    |||
93315 ACAGTACTAAAGTCTTACCGACAGACATCAACAAAGAAAGAAATGAAGGGATTCOA 93374
    |||
539 ATCAATPAAGAGAAAGTCAAACTGTCCTGTTCACTGATGATGATGTTATACCTAGA 598
    |||
93375 ATTGTTAAAGAGAAAGTCAATCTGTCACGTTTGTGATGATGATGTTATACCTCAA 93434
    |||
599 AACCTTAAAGACTCATCCAGAAAGCTCCTGAACTGATACATAAATTCAGTAAGTTCA 658
    |||
93435 AACCTTAAAGACTCCTCTAGAAAGCTCCTGAACTGATPAAGATTCACAAAGTTTCC 93494
    |||
659 GGATCAAACTAAATGTACAAAAATCAGTACGCTATACAGCAAGTACCAAGCT 718
    |||
93495 AGATACAAATTAATATATACAAATTCAGTCTTTTAAACCAACACCAACCAAGCA 93554
    |||
719 GGAATCAATCAATCAAGAACTCAAACTTTTACATAGCTGTAAAAA-----TACTT 771
    |||
93555 GGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 93614
    |||
772 AAGATATTTCTTACCCAGAGGTGAGAGACCTCTTCAAGGAAATCTCAAAACCACT 831
    |||
93615 AAGATATTTCTTACCCAGAGGTGAGAGACCTCTTCAAGGAAATCTCAAAACCACT 93674
    |||
832 ----GACATCATAGATGACAAACAAAGTGAACAACTCCATGCTCATGAGGATG 887
    |||
93675 GAAAGAAATCATAGATGACAAACAAAGTGAACAACTCCATGCTCATGAGGATG 93734
    |||
888 AATCAATATTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 947
    |||
93735 AATCAATATTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93794
    |||
948 CACCAAAATATCATCTATCTTCTTCAAGAACTTGAAGAAATCTTGAATTCAT 1007
    |||
93795 CATCAAAACCAACCATCTATCTTCTTCAAGAACTTGAAGAAATCTTGAATTCAT 93853
    |||
1008 GGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1067
    |||
93854 GGAAC-----CAAAAGAGTCCATGACCAAGCAAGCAAGCAAGCAAGCAAGCA 93898
    |||
1068 GAACAATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1127
    |||
93899 GAACAATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93958
    |||
1128 CAAAACATCTGACATGACATTAATGAGCATGACATGACATGAGAAAGAGAGAA 1187
    |||
93959 CAAAACATCTGATCTGATTAATTAATGAGCATGACATGACATGAGAAAGAGAGAA 94018
    |||
1188 TCAGAAATTAAGCCAAATTAATTAATGAGCAATGATTTTGAAGAAAGCAAAACAT 1247
    |||
94019 CCGAAGAAATTAAGCCAAACATTAAGCAATGATCTTTCGAGAAAGCAAAACAT 94078
    |||
1248 AAGTGGGAAAGACATTTCTAGTTAACAATGATGATGATGATGATGATGATGATGAT 1307
    |||
94079 AAGTGGGAAAGACATTTCTTAACAATATGATGATGATGATGATGATGATGATGAT 94138
    |||
1308 TGAAGATGAAGCTGATCCCTGCTCTTCACTTAATCAAAATTTGATACAAAGTGA 1367
    |||
94139 TGAAGATGAAGCTGAT-CTCATCTCTCACTTATCAAAATTCACATCAAGATGA 94197
    |||

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 1, 2004, 23:06:03; Search time 3465 Seconds
(without alignments)
13062.196 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622
Sequence: 1 tgcctgaagccatgcactc.....gtccccaagcccttacc 8622

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23sep04:*
1: Geneseq19808:*
2: Geneseq19908:*
3: Geneseq20008:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8622	100.0	8622	6	ABK52823
2	1401	16.2	66494	6	ABO88140 Human oet
3	1390.4	16.1	110000	6	ABR90193 Human oes
4	1380.4	16.1	110000	6	ABQ87681 Human oes
5	1390.4	16.0	110000	8	ABX33717 Gene enco
6	1375.6	16.0	110000	6	ABX08336.13
7	1375.6	16.0	110000	12	ADJ25985.13
8	1375.6	16.0	110000	6	ADN97989_13
9	1374.6	15.9	1400	6	ADJ33751
10	1374.6	15.9	1400	8	ABZ42595
11	1374.6	15.9	1400	12	ADH08534
12	1374	15.8	110000	12	ADDO50281_13
13	1358	15.8	99014	6	ABN96931
14	1355	15.7	42738	4	AAK68992
15	1346	15.6	136284	6	ABK83575
16	1311.6	15.2	110000	4	AAK95240_05
17	1311.6	15.2	110000	4	AAK96733_05
18	1311.6	15.2	110000	6	ABT00010_05
19	1311.6	15.2	110000	6	ABT01503_05
20	1311.6	15.2	110000	12	ADH77485_05
21	1308	15.2	349881	10	ADC66642

C	22	1302	15.1	145606	12	ADP44617
	23	1301.8	15.1	32167	4	AAI05509
	24	1288.6	14.9	247682	12	ADL08109
	25	1286.8	14.9	148567	9	ACA62841
	26	1286.8	14.9	148567	10	ABR55500
	27	1286.8	14.9	148567	12	ABR55500
	28	1280.4	14.9	70665	6	ABT10716
	29	1280.4	14.9	70665	11	ADN95672
	30	1274	14.8	1369	10	ADC68820
	31	1271.6	14.7	32249	4	AAI05336
	32	1271.6	14.7	32249	4	ABL98205
	33	1264	14.7	1604	6	ABD33752
	34	1264	14.7	1604	8	ABZ42602
	35	1264	14.7	1604	12	ADH08536
	36	1247	14.5	79590	9	ADL27152
	37	1247	14.5	79590	9	ADA03074
	38	1247	14.5	79684	9	ADA66358
	39	1247	14.5	79684	10	ADH72812
C	40	1245.6	14.4	136284	6	ABK83575
	41	1243.8	14.4	4019	10	AAK99601
	42	1228.4	14.2	2040	6	ADH08519
	43	1228.4	14.2	2040	12	ADH08519
	44	1228.4	14.2	2040	12	ADH08519
C	45	1225.8	14.2	110000	12	ADH69807_0

ALIGNMENTS

RESULT 1	ABK52823	standard; DNA; 8622 BP.
ID	ABK52823	standard; DNA; 8622 BP.
XX	ABK52823	
AC	ABK52823	
XX	27-AUG-2002	(first entry)
DT	27-AUG-2002	(first entry)
XX	Genomic DNA encoding human G-protein coupled receptor (GPCR).	
XX	Human; G-protein coupled; receptor; GPCR; human protease;	
KW	human therapeutic protein; query sequence; search; gene; ds;	
KW	sequence database; non-human transgenic animal; gene therapy;	
KW	Chromosome 3.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	4300..8496
FT		/*tag= a
FT		/product= "Human G-protein coupled receptor (GPCR) "
FT	exon	4300..4319
FT		/*tag= b
FT	intron	4320..7502
FT		/*tag= c
FT	exon	/number= 1
FT		/number= 1
FT		7503..8496
FT		/*tag= d
FT		/number= 2
XX	WO200234914-A1.	
XX		
XX	02-MAY-2002.	
PD	02-MAY-2002.	
XX		
PF	10-OCT-2001; 2001WO-US031592.	
XX		
PR	25-OCT-2000; 2000US-00695045.	
XX		
PR	31-MAY-2001; 2001US-00867570.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM,	
XX		

DR WPI: 2002-463360/49.
DR P-PSDB; AAU97598.
XX Novel isolated G-protein coupled receptor peptide useful for treating
PT disorder characterized by absence of, in appropriate or unwanted
PT expression of the receptor protein, and as immunogens to raise
PT antibodies.
XX
XX Claim 4; Fig 3; 75bp; English.
XX
XX The present invention relates to a new G-protein coupled receptor (GPCR)
XX peptide. The invention is useful for identifying a modulator of GPCR and
XX for treating a disease or condition mediated by a human protease. The
XX invention is also useful as models for the development of human
XX therapeutics, for identifying therapeutic proteins, as targets for
XX development of human therapeutic agents, and as query sequence to perform
XX a search against sequence databases to, for e.g., identify other family
XX members of related sequences. The vector of the invention is useful for
XX producing a GPCR protein or peptide, for conducting cell-based assays
XX involving the GPCR protein or its fragment, for identifying GPCR protein
XX mutants whose functions are affected, and to produce non-human transgenic
XX animals. The present nucleic acid sequence represents the human G-protein
XX coupled receptor (GPCR) gene located on chromosome 3. This sequence
XX encodes the human G-protein coupled receptor (GPCR) protein of the
XX invention
SQ Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other;
Query Match 100.0%; Score 8622; DB 6; Length 8622;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTATGAAGCCATGCTCTTTATATACCAAAACAGAGAAAGATTTTCAAAAAGAAAC 60
DB 1 TGTATGAAGCCATGCTCTTTATATACCAAAACAGAGAAAGATTTTCAAAAAGAAAC 60
QY 61 TATAGACCAAGTACCACTGATGATATATATATATATATATATATATATATATATAT 120
DB 61 TATAGACCAAGTACCACTGATGATATATATATATATATATATATATATATATATAT 120
QY 121 CCAATCCACAGCATATCAAGAAAGATATATATATATATATATATATATATATATATAT 180
DB 121 CCAATCCACAGCATATCAAGAAAGATATATATATATATATATATATATATATATATAT 180
QY 181 GTGCGAGATAGTTAT 240
DB 181 GTGCGAGATAGTTAT 240
QY 241 AAAACAAAAT 300
DB 241 AAAACAAAAT 300
QY 301 ATTTCTTTAT 360
DB 301 ATTTCTTTAT 360
QY 361 TAAAGGCAAT 420
DB 361 TAAAGGCAAT 420
QY 421 CATTTCTCTGAGAACTGGAACAAGATGCTATCTTCAACCTTCTATTTCAACAT 480
DB 421 CATTTCTCTGAGAACTGGAACAAGATGCTATCTTCAACCTTCTATTTCAACAT 480
QY 481 AGTAGTGAAGTTTATGCGAGCAATCAAGAAAGAAAGAAATCAAGGGCAACCAAT 540
DB 481 AGTAGTGAAGTTTATGCGAGCAATCAAGAAAGAAAGAAATCAAGGGCAACCAAT 540
QY 541 CAATAAAGGAGAGTAACTGCTCTGCTTCACTGATGATATATATATATATATATATAT 600
DB 541 CAATAAAGGAGAGTAACTGCTCTGCTTCACTGATGATATATATATATATATATATAT 600
QY 601 CCTTAAAGACTATCCAGAAAGCTCTAGAACTGATATATATATATATATATATATATAT 660

DB 601 CCTTAAAGACTATCCAGAAAGCTCTAGAACTGATATATATATATATATATATATATAT 660
QY 661 ATACAAAT 720
DB 661 ATACAAAT 720
QY 721 GATCAAAATCAAGAACTGAAACCTTTTCAATAGCTGTAATATATATATATATATATAT 780
DB 721 GATCAAAATCAAGAACTGAAACCTTTTCAATAGCTGTAATATATATATATATATATAT 780
QY 781 CTATCCAGAGAGTGAAGAGCTCTACAGAGAAATCTACAAACAGCTGACATCTATA 840
DB 781 CTATCCAGAGAGTGAAGAGCTCTACAGAGAAATCTACAAACAGCTGACATCTATA 840
QY 841 GATGACAAACAAGTGAAGCAATCCATCTCATGATGATGATGATGATGATGATGATG 900
DB 841 GATGACAAACAAGTGAAGCAATCCATCTCATGATGATGATGATGATGATGATGATG 900
QY 901 AAAATGACATATTTGCAAGCAATCTTCAAGTTCATGCAATTCACCAACCAATATCA 960
DB 901 AAAATGACATATTTGCAAGCAATCTTCAAGTTCATGCAATTCACCAACCAATATCA 960
QY 961 TCATCTTTCTGACAGAACTAGAAAACATTTCTTAAATTCATATGGAACAACACA 1020
DB 961 TCATCTTTCTGACAGAACTAGAAAACATTTCTTAAATTCATATGGAACAACACA 1020
QY 1021 AAAAAAAAAAAAAAAAACCCGCATAGCAAGCAAGCACTTACGAAAGCAAAATCTGCA 1080
DB 1021 AAAAAAAAAAAAAAAAACCCGCATAGCAAGCAAGCACTTACGAAAGCAAAATCTGCA 1080
QY 1081 GGCATCACATTTCCCATCTTCAACATATCTCAAGGCTATATATATATATATATATATAT 1140
DB 1081 GGCATCACATTTCCCATCTTCAACATATCTCAAGGCTATATATATATATATATATATAT 1140
QY 1141 CATGACATTAACCTAGGACATATGCAATGGAAGAAAGAGAAATCCAGAAATTAAG 1200
DB 1141 CATGACATTAACCTAGGACATATGCAATGGAAGAAAGAGAAATCCAGAAATTAAG 1200
QY 1201 CCAAT 1260
DB 1201 CCAAT 1260
QY 1261 GACATTTCTAT 1320
DB 1261 GACATTTCTAT 1320
QY 1321 CTGATCCCTTGTCTCTCACTTATATCAAAAATTTGATCAAGATGATCAAAAGCTTAA 1380
DB 1321 CTGATCCCTTGTCTCTCACTTATATCAAAAATTTGATCAAGATGATCAAAAGCTTAA 1380
QY 1381 TCTGAGACTTAAACCAT 1440
DB 1381 TCTGAGACTTAAACCAT 1440
QY 1441 TCACCTAGGCAAGAACTTATGAGCAAGAAACCAAAAGTAAATGCAACAAAACAAAT 1500
DB 1441 TCACCTAGGCAAGAACTTATGAGCAAGAAACCAAAAGTAAATGCAACAAAACAAAT 1500
QY 1501 AATATGATGAGCTTAT 1560
DB 1501 AATATGATGAGCTTAT 1560
QY 1561 AGCAAAACAACAACCAAGAGTGAAGAAATCTTCAAACTAAGATCTGATTAAGG 1620
DB 1561 AGCAAAACAACAACCAAGAGTGAAGAAATCTTCAAACTAAGATCTGATTAAGG 1620
QY 1621 ACTAATATCCGAATCAAGAACTCAAACTAAGATCTGATTAAGG 1680
DB 1621 ACTAATATCCGAATCAAGAACTCAAACTAAGATCTGATTAAGG 1680
QY 1681 CATGAAAGAGTGGCTTAAGACATGATATGCAATTTCTCAAAAGAAAGATATATCAATG 1740

Dh 1681 CATGAAGATGGGCTAAGACATGATGACAAATTCCTCAAAAGAAAGATATACAAATGGC 1740
Qy 1741 CAAGAAACAGGAAAAATGCTTAAACATGATGATTTAGGAAAATGTAATCAACACTG 1800
Db 1741 CAAGAAACAGGAAAAATGCTTAAACATGATGATTTAGGAAAATGTAATCAACACTG 1800
Qy 1801 TTAATGCGATACACCTTATCTCTGCAAGATGTCATATTTAAATCTAAATTAATA 1860
Db 1801 TTAATGCGATACACCTTATCTCTGCAAGATGTCATATTTAAATCTAAATTAATA 1860
Qy 1861 GATGTGGTGGGTCTGTGTGATTAAGAGACATTTTACCTGCTGTGGATGTAAAC 1920
Db 1861 GATGTGGTGGGTCTGTGTGATTAAGAGACATTTTACCTGCTGTGGATGTAAAC 1920
Qy 1921 TTGGGCAACCACTATGAAAAAGGTGGAAATTTCTTAAGAACTAAAGTAGATGAC 1980
Db 1921 TTGGGCAACCACTATGAAAAAGGTGGAAATTTCTTAAGAACTAAAGTAGATGAC 1980
Qy 1981 CATTTGATCCAGCAATCCCATTAATAATGTAATAATATATTTATATATCATGGAAT 2040
Db 1981 CATTTGATCCAGCAATCCCATTAATAATGTAATAATATATTTATATATCATGGAAT 2040
Qy 2041 ACAACTGAGCCATATAAAAGATTAATGATGACATTCACAGCAATCTAGATGGAAATGG 2100
Db 2041 ACAACTGAGCCATATAAAAGATTAATGATGACATTCACAGCAATCTAGATGGAAATGG 2100
Qy 2101 AGACCCCTTATTTCTAAGTGGGGTAACTCAGAAATGAAAAACCAACATCATATGTTCTTAC 2160
Db 2101 AGACCCCTTATTTCTAAGTGGGGTAACTCAGAAATGAAAAACCAACATCATATGTTCTTAC 2160
Qy 2161 TTAAAGTGGGGGCTAAGCTGTGAGAGACAGAGGCAATGAAATGATATATATGAACTCTGG 2220
Db 2161 TTAAAGTGGGGGCTAAGCTGTGAGAGACAGAGGCAATGAAATGATATATATGAACTCTGG 2220
Qy 2221 GGACTTGAAGGGAAGATGAAAGAGAGGCGAGGATAAAGACTACACATGGGTACAGT 2280
Db 2221 GGACTTGAAGGGAAGATGAAAGAGAGGCGAGGATAAAGACTACACATGGGTACAGT 2280
Qy 2281 GTACACTGCTCAGGTGATGGGTGCAACAAATCTCAGAAATTAACAATAAGAACTTATC 2340
Db 2281 GTACACTGCTCAGGTGATGGGTGCAACAAATCTCAGAAATTAACAATAAGAACTTATC 2340
Qy 2341 CATGGAAGCAACACCCCTGTTCCCAAAATCCCAATGAATTAATAATTAATAATA 2400
Db 2341 CATGGAAGCAACACCCCTGTTCCCAAAATCCCAATGAATTAATAATTAATAATA 2400
Qy 2401 ATGATTTAATTTCAAGAAATTTAAAAAGTCACTGTTCAAGTTTATATATATGAAGTA 2460
Db 2401 ATGATTTAATTTCAAGAAATTTAAAAAGTCACTGTTCAAGTTTATATATATGAAGTA 2460
Qy 2461 AGAATGAAAAAGTGTAGCAAGTGTAGCCTGTGCAATGGAATCTAGATTTTCACTTGG 2520
Db 2461 AGAATGAAAAAGTGTAGCAAGTGTAGCCTGTGCAATGGAATCTAGATTTTCACTTGG 2520
Qy 2521 CATACACTTCTCTGCAATTTGAAAAGAAAGTATACATGAAATATATACACATATGATA 2580
Db 2521 CATACACTTCTCTGCAATTTGAAAAGAAAGTATACATGAAATATATACACATATGATA 2580
Qy 2581 AAGAAAAATCATCAAAAAATGTGTGACAGGCATTTGTCAAGCTTGAAATGGTCCCATGATCTAC 2640
Db 2581 AAGAAAAATCATCAAAAAATGTGTGACAGGCATTTGTCAAGCTTGAAATGGTCCCATGATCTAC 2640
Qy 2641 TTTTTCATTTGGATATAAAGCCTCATATATGATGATTCATTTGCTTAATGTATGCTTAC 2700
Db 2641 TTTTTCATTTGGATATAAAGCCTCATATATGATGATTCATTTGCTTAATGTATGCTTAC 2700
Qy 2701 GCCCATATTTGATTTTAAATCAGAGACAAATTAATTAACAGAAAGTTGAACAAGATGG 2760
Db 2701 GCCCATATTTGATTTTAAATCAGAGACAAATTAATTAACAGAAAGTTGAACAAGATGG 2760
Qy 2761 GACGTGATAGAGAGGCTTAATATGATGATGATGAGACAGAGGCAAGATCATCTCAG 2820
Db 2761 GACGTGATAGAGAGGCTTAATATGATGATGATGAGACAGAGGCAAGATCATCTCAG 2820

Qy 2821 TTAGGATTTGTGTCTCAAAATCTCTGGCCTCTGATATTTGCCCATAGTCTCATACAGAA 2880
Db 2821 TTAGGATTTGTGTCTCAAAATCTCTGGCCTCTGATATTTGCCCATAGTCTCATACAGAA 2880
Qy 2881 ATTAACAGATGTCACACATCTTGTGAAGCCTGGAATGTCACAGAGTTTCAATTCAGCT 2940
Db 2881 ATTAACAGATGTCACACATCTTGTGAAGCCTGGAATGTCACAGAGTTTCAATTCAGCT 2940
Qy 2941 CCTGTAGGCACTCTCTGAATTAAGCAACAGAAAGTCTCTGAAGTCACTGAATCCCA 3000
Db 2941 CCTGTAGGCACTCTCTGAATTAAGCAACAGAAAGTCTCTGAAGTCACTGAATCCCA 3000
Qy 3001 GAAAGGCTCTTACCTTTAGCAAGGAGGTCTTACCACTGCAACAAAGAAAGAGAT 3060
Db 3001 GAAAGGCTCTTACCTTTAGCAAGGAGGTCTTACCACTGCAACAAAGAAAGAGAT 3060
Qy 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGTCAAGTATGATTTTGTGTAAGATC 3120
Db 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGTCAAGTATGATTTTGTGTAAGATC 3120
Qy 3121 ATGTCTTATGCTTCCACCTTGTGTCTACATGACAGGGGTTCAGAGCTTGTTCAGGAAA 3180
Db 3121 ATGTCTTATGCTTCCACCTTGTGTCTACATGACAGGGGTTCAGAGCTTGTTCAGGAAA 3180
Qy 3181 AGACAGAGACATGAACTTCTTTCAGAAACTGAGTGTCTCAACCCAACTGTGTAG 3240
Db 3181 AGACAGAGACATGAACTTCTTTCAGAAACTGAGTGTCTCAACCCAACTGTGTAG 3240
Qy 3241 CTCTAAATGATGATCCCTCTTAAATTAATCTCCCATATCACTCTCTCAATTCATCA 3300
Db 3241 CTCTAAATGATGATCCCTCTTAAATTAATCTCCCATATCACTCTCTCAATTCATCA 3300
Qy 3301 TTCAATCTGCTCATGAGAGAGACTGTGCTCTTATCATTTATTTAAGCAAGAGGAC 3360
Db 3301 TTCAATCTGCTCATGAGAGAGACTGTGCTCTTATCATTTATTTAAGCAAGAGGAC 3360
Qy 3361 ATGCAAGCAATTTCTTCCAGAGTGAACCTGTATAGAGCAGATTCCTTGTGTTCACTTAC 3420
Db 3361 ATGCAAGCAATTTCTTCCAGAGTGAACCTGTATAGAGCAGATTCCTTGTGTTCACTTAC 3420
Qy 3421 TTTTCAATTTATCTCTTGTGCTATCTGAAAGTCTAAGAAATATATAGAGGCCCA 3480
Db 3421 TTTTCAATTTATCTCTTGTGCTATCTGAAAGTCTAAGAAATATATAGAGGCCCA 3480
Qy 3481 ATTAATTAAGAGTGTTCATGAGAAAGTATTTAACAAGATGACAGAGTTAAGGCTCAG 3540
Db 3481 ATTAATTAAGAGTGTTCATGAGAAAGTATTTAACAAGATGACAGAGTTAAGGCTCAG 3540
Qy 3541 GATCTTAAGCAATACATATAGGGGACACTTCTCTCCCTTGAAGCTGAAACGAGACG 3600
Db 3541 GATCTTAAGCAATACATATAGGGGACACTTCTCTCCCTTGAAGCTGAAACGAGACG 3600
Qy 3601 GAAGAGCAGTTACATTTGTGCGCATAGCCATAGCTGTAGCCATTAAGGTTGGAGAGCAT 3660
Db 3601 GAAGAGCAGTTACATTTGTGCGCATAGCCATAGCTGTAGCCATTAAGGTTGGAGAGCAT 3660
Qy 3661 GAGAGGCAAGTGAAGAAAGCCTGCTGCGCAACAGCAGCAGCAGCAGGCTGATATAGTT 3720
Db 3661 GAGAGGCAAGTGAAGAAAGCCTGCTGCGCAACAGCAGCAGCAGCAGGCTGATATAGTT 3720
Qy 3721 TGATCTGTGTCTCCACAAATCTCATGTTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
Db 3721 TGATCTGTGTCTCCACAAATCTCATGTTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
Qy 3781 CCTGTGAGAGATATTAATGATCAGGGGAGATGTTTTCATGAAATGTTTAAACACAT 3840
Db 3781 CCTGTGAGAGATATTAATGATCAGGGGAGATGTTTTCATGAAATGTTTAAACACAT 3840
Qy 3841 CCCCTTTGATATGTTGTGTGATCTGACAGAGTCTCATGAAATCTAGTTGTTTAAA 3900
Db 3841 CCCCTTTGATATGTTGTGTGATCTGACAGAGTCTCATGAAATCTAGTTGTTTAAA 3900

QY 3901 GCGTGTAGACACCTCCTCCTCTCTCTTACTCTGCTCTACAGATGTAGAGCGCTGCTC 3960
Db 3901 GCGTGTAGACACCTCCTCCTCTCTCTTACTCTGCTCTACAGATGTAGAGCGCTGCTC 3960
QY 3961 CCCCTTTTGGCTTTTACACAGAGATTGGAAGCTTCTGAGGCTCTCCAGAGCAGAGCTGC 4020
Db 3961 CCCCTTTTGGCTTTTACACAGAGATTGGAAGCTTCTGAGGCTCTCCAGAGCAGAGCTGC 4020
QY 4021 TATGCTCTCTGTATGAGTCTGTAGAGCTATTAGGCACTTAAACCATTTCTCTATTAAT 4080
Db 4021 TATGCTCTCTGTATGAGTCTGTAGAGCTATTAGGCACTTAAACCATTTCTCTATTAAT 4080
QY 4081 TCCCACTCTCAGGATTTCTTTTATGCAATTGAGATGAACTAATACAGACAGAGAG 4140
Db 4081 TCCCACTCTCAGGATTTCTTTTATGCAATTGAGATGAACTAATACAGACAGAGAG 4140
QY 4141 CCAGAGATGAAATCCCAAGGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
Db 4141 CCAGAGATGAAATCCCAAGGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
QY 4201 CCAGGTCTCAATTCACACAGAAACAGAAATTAAGAAATCCCACTGATGTGATATA 4260
Db 4201 CCAGGTCTCAATTCACACAGAAACAGAAATTAAGAAATCCCACTGATGTGATATA 4260
QY 4261 GAAGCACTCTCTGGGATGTCAACAGATTAAGAAAGATGGAAGCAATCTCATG 4320
Db 4261 GAAGCACTCTCTGGGATGTCAACAGATTAAGAAAGATGGAAGCAATCTCATG 4320
QY 4321 TAAATGAGATCTCTCTCTCACTTCTGTATCTCTCTAAATCTCTGGGCTTCTCTATC 4380
Db 4321 TAAATGAGATCTCTCTCTCACTTCTGTATCTCTCTAAATCTCTGGGCTTCTCTATC 4380
QY 4381 TGATGATCCGCTCTCATTTGAGCTCTATCAGACTTAATGTTGGCTTGTCTTC 4440
Db 4381 TGATGATCCGCTCTCATTTGAGCTCTATCAGACTTAATGTTGGCTTGTCTTC 4440
QY 4441 TCTACTGTCACTTTATGCAAGAAATGTGCACTTTGTTAAATGCAATAGAAATAAAT 4500
Db 4441 TCTACTGTCACTTTATGCAAGAAATGTGCACTTTGTTAAATGCAATAGAAATAAAT 4500
QY 4501 GTTAATTTTAAAGAAATATGATATTTTGTGTAATATTAAGTTGGCTGATCTAATAA 4560
Db 4501 GTTAATTTTAAAGAAATATGATATTTTGTGTAATATTAAGTTGGCTGATCTAATAA 4560
QY 4561 GACATGAGAGAAATATCTTAAACAGAAAGTATGTGCTCTGGGCTCACTAGGT 4620
Db 4561 GACATGAGAGAAATATCTTAAACAGAAAGTATGTGCTCTGGGCTCACTAGGT 4620
QY 4621 CTGAATCTACAGATTCACAACTACAGAGAACTTTTCCAAAATTAAGGTGTGCG 4680
Db 4621 CTGAATCTACAGATTCACAACTACAGAGAACTTTTCCAAAATTAAGGTGTGCG 4680
QY 4681 GAGTTGTATGTATCTGAACAGGTACAACTTGTATTTCTTGTCAATATTTCTGAAAA 4740
Db 4681 GAGTTGTATGTATCTGAACAGGTACAACTTGTATTTCTTGTCAATATTTCTGAAAA 4740
QY 4741 CTACATATTAACAGAACTTATATAGATTTGCAATTTGTCAATATTTCTAATCTT 4800
Db 4741 CTACATATTAACAGAACTTATATAGATTTGCAATTTGTCAATATTTCTAATCTT 4800
QY 4801 AAATGATTTATGTATCTGGAGAAAGTGTAGATATATACAAATACATATATAGG 4860
Db 4801 AAATGATTTATGTATCTGGAGAAAGTGTAGATATATACAAATACATATATAGG 4860
QY 4861 AAATGAGATCTGAGATTTTGTGTGTGTGGGCTTCTGAAAGAAATCCCTGTAAAT 4920
Db 4861 AAATGAGATCTGAGATTTTGTGTGTGTGGGCTTCTGAAAGAAATCCCTGTAAAT 4920
QY 4921 ACACAAAAATGACACTCTGAGATCTGAATAGAGCTCCAAAGCATCATACATGAG 4980
Db 4921 ACACAAAAATGACACTCTGAGATCTGAATAGAGCTCCAAAGCATCATACATGAG 4980
QY 4981 TTCAAAAATTTGCTGCTCCCAAGTCTTAGAGAGTGGCTCATCTTGTATCTAAT 5040

Db 4981 TTCAAAAATTTGCTGCTCCCAAGTCTTAGAGAGTGGCTCATCTTGTATCTAAT 5040
QY 5041 GGTTCACAGCATTTAGATTCAGCTTTATAGAAAAAGACAGGGGAAAGAGCT 5100
Db 5041 GGTTCACAGCATTTAGATTCAGCTTTATAGAAAAAGACAGGGGAAAGAGCT 5100
QY 5101 TTGCTCTCTTATTAATCCCATGAGCAGGACTTGTCTGTGCACTTTTGTATCTTCC 5160
Db 5101 TTGCTCTCTTATTAATCCCATGAGCAGGACTTGTCTGTGCACTTTTGTATCTTCC 5160
QY 5161 ACTTAAACAGACCTGCTCAATGGATGTCAATCAGATCAAGAAACTGGATGTGCTC 5220
Db 5161 ACTTAAACAGACCTGCTCAATGGATGTCAATCAGATCAAGAAACTGGATGTGCTC 5220
QY 5221 CTGTGCTCTGTATCAATCTCAGAAAGTTATGTGACCAAAAAAGAAATCTTGGGCA 5280
Db 5221 CTGTGCTCTGTATCAATCTCAGAAAGTTATGTGACCAAAAAAGAAATCTTGGGCA 5280
QY 5281 ACCAGAGCTCTTCAAGCCCTGAGCTGTCTGATTTGTGTGCTCATCAAGATTTTCA 5340
Db 5281 ACCAGAGCTCTTCAAGCCCTGAGCTGTCTGATTTGTGTGCTCATCAAGATTTTCA 5340
QY 5341 GGAATCTCTCAGAAATTAATAAATGTGTGGGACAGAACTGAGTCTTGTGACAGA 5400
Db 5341 GGAATCTCTCAGAAATTAATAAATGTGTGGGACAGAACTGAGTCTTGTGACAGA 5400
QY 5401 CTCACAGGACAGGGGCTGTATGTGACCTGCTCTTAATGTTGTGAAACAGAAAACT 5460
Db 5401 CTCACAGGACAGGGGCTGTATGTGACCTGCTCTTAATGTTGTGAAACAGAAAACT 5460
QY 5461 TTAATCTCTAAGCCCTTACCTCATCTTAATGTTATAGATTAATACATAGACGTCT 5520
Db 5461 TTAATCTCTAAGCCCTTACCTCATCTTAATGTTATAGATTAATACATAGACGTCT 5520
QY 5521 TTAAGAAATCAATAGCATGTTAAACACATCTAAATGTGTGATACACAGTAA 5580
Db 5521 TTAAGAAATCAATAGCATGTTAAACACATCTAAATGTGTGATACACAGTAA 5580
QY 5581 AGACAGGATGACTTACTCTCTTAAGGATCTTGGGTTTCAATGAGAAACAACTATCA 5640
Db 5581 AGACAGGATGACTTACTCTCTTAAGGATCTTGGGTTTCAATGAGAAACAACTATCA 5640
QY 5641 TACCATACCTAATAGAGGACAAACAGTTAGTGTCTGAGTGTGATTAACAGAGTTCTC 5700
Db 5641 TACCATACCTAATAGAGGACAAACAGTTAGTGTGTCTGAGTGTGATTAACAGAGTTCTC 5700
QY 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGAGCTTGTCTCTAAG 5760
Db 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGAGCTTGTCTCTAAG 5760
QY 5761 AGCTATGATGATGACACTCTCTGATGCTCTTAACTCCAGAGAGGAGATGATC 5820
Db 5761 AGCTATGATGATGACACTCTCTGATGCTCTTAACTCCAGAGAGGAGATGATC 5820
QY 5821 TTCTTCACTGAGAGCTCTGCGCATGTGCTTGAATCTTGGATTCAGAGAACTGCG 5880
Db 5821 TTCTTCACTGAGAGCTCTGCGCATGTGCTTGAATCTTGGATTCAGAGAACTGCG 5880
QY 5881 CCTGTATAGGCAAAAAAGAACTCTGAATTTGTTCTTAATGAGCACTCACTCTA 5940
Db 5881 CCTGTATAGGCAAAAAAGAACTCTGAATTTGTTCTTAATGAGCACTCACTCTA 5940
QY 5941 TTTTCCCTTATTTCAATTTGCTTCTCATCTCATCTGAGTTGTATAGTTATTTT 6000
Db 5941 TTTTCCCTTATTTCAATTTGCTTCTCATCTCATCTGAGTTGTATAGTTATTTT 6000
QY 6001 TTTTCAAGCCCAATTTTGAATGATCTGATTAATCTGAGATCACTCTCTACT 6060
Db 6001 TTTTCAAGCCCAATTTTGAATGATCTGATTAATCTGAGATCACTCTCTACT 6060
QY 6061 TTAACCCCTCTAATGATTAATGACATAGTGTGCTGCTGCAAGGCTGTGCA 6120

Db 6061 TTACCCCCCTCTAATCATGTATATGACACATAGTGTGCTGGTCCAAAGGCGTGTGA 6120
 Qy 6121 AAAAATGATCATGATGACGCTTGTGGGCTCACTTCACTATATATATATATACCC 6180
 Db 6121 AAAAATGATCATGATGACGCTTGTGGGCTCACTTCACTATATATATATATACCC 6180
 Qy 6181 CCGGGAAGAAATATATCTCTCTTATATCTATATTTCAATATGCGAATACAGAAATTA 6240
 Db 6181 CCGGGAAGAAATATATCTCTCTTATATCTATATTTCAATATGCGAATACAGAAATTA 6240
 Qy 6241 TACTAATTTTACCTCTAGGCTCTTGCATGATTAAGAGGCAATACCTATTAATACG 6300
 Db 6241 TACTAATTTTACCTCTAGGCTCTTGCATGATTAAGAGGCAATACCTATTAATACG 6300
 Qy 6301 TCAATCAGCTCTGTATATCTCCCAATTAGACCTATCTCATTTCTCAGTTGAATTT 6360
 Db 6301 TCAATCAGCTCTGTATATCTCCCAATTAGACCTATCTCATTTCTCAGTTGAATTT 6360
 Qy 6361 GCATGAATATCTCTCTTACAAACCAAGCCCTACACTTCTCTATTTCTCACTGAGCT 6420
 Db 6361 GCATGAATATCTCTCTTACAAACCAAGCCCTACACTTCTCTATTTCTCACTGAGCT 6420
 Qy 6421 CCTCTCATCAAAATGTTTGCATCAACAAAGAAAGGCTACAAAGATCTCCGAAAGAG 6480
 Db 6421 CCTCTCATCAAAATGTTTGCATCAACAAAGAAAGGCTACAAAGATCTCCGAAAGAG 6480
 Qy 6481 AATGAATAGGTTTACATTTGTATCTCAGCAGAAACCTTAGTAGTCCCATATCATAT 6540
 Db 6481 AATGAATAGGTTTACATTTGTATCTCAGCAGAAACCTTAGTAGTCCCATATCATAT 6540
 Qy 6541 TCCCACTTCAATTAATCTGTGCTGAGTGGCACTCAGGCTCACCTCATTTCTCTTCT 6600
 Db 6541 TCCCACTTCAATTAATCTGTGCTGAGTGGCACTCAGGCTCACCTCATTTCTCTTCT 6600
 Qy 6601 CTGTTCTATGCTGAGCAATCAGCTCAGACCAACCTTACCCAAACATCTGTGACAA 6660
 Db 6601 CTGTTCTATGCTGAGCAATCAGCTCAGACCAACCTTACCCAAACATCTGTGACAA 6660
 Qy 6661 ATGCTTCAAGGGTTCGGCAAGCCCACTGAGTCTTATTTTAAAGGCATCATGAGGT 6720
 Db 6661 ATGCTTCAAGGGTTCGGCAAGCCCACTGAGTCTTATTTTAAAGGCATCATGAGGT 6720
 Qy 6721 CAATTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACACAGATGAGCTGTCCA 6780
 Db 6721 CAATTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACACAGATGAGCTGTCCA 6780
 Qy 6781 AATGAATCTGATGAATGAATTTTCTCATCTAATATATATGATGATCTTATGCT 6840
 Db 6781 AATGAATCTGATGAATGAATTTTCTCATCTAATATATATGATGATCTTATGCT 6840
 Qy 6841 TACATTTGCTTCAATTTTATTTTCAATTTTCAATCAATCTACATTTGCCATTAGGCT 6900
 Db 6841 TACATTTGCTTCAATTTTATTTTCAATTTTCAATCAATCTACATTTGCCATTAGGCT 6900
 Qy 6901 CTGATGATGATCTTCTCATTTGAAATGAACGTTTATGAAAAGCAATTTGCTGCTTATG 6960
 Db 6901 CTGATGATGATCTTCTCATTTGAAATGAACGTTTATGAAAAGCAATTTGCTGCTTATG 6960
 Qy 6961 GAATAGGCACTAGAGATATAAATGTAAATGTGATGCTGCTGCTGCAATGATGACACT 7020
 Db 6961 GAATAGGCACTAGAGATATAAATGTAAATGTGATGCTGCTGCTGCAATGATGACACT 7020
 Qy 7021 GAGTTATTTCTCACCCACAGGCTCCGCAATTTTCAACATCTTAGGAAATCCCAATTT 7080
 Db 7021 GAGTTATTTCTCACCCACAGGCTCCGCAATTTTCAACATCTTAGGAAATCCCAATTT 7080
 Qy 7081 TCCCTGCTTCAATATGATGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7140
 Db 7081 TCCCTGCTTCAATATGATGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7140
 Qy 7141 AGGCTGTCAACAACCTCTTGTGTATCTGAATTTCTTCCACCTGAGAGAAATTTTCAG 7200
 Db 7141 AGGCTGTCAACAACCTCTTGTGTATCTGAATTTCTTCCACCTGAGAGAAATTTTCAG 7200

Qy 7201 CCGAGATAGATATCATCGGCTTCCACAGACTGCTAGATGATGAGGCTGTTTAT 7260
 Db 7201 CCGAGATAGATATCATCGGCTTCCACAGACTGCTAGATGATGAGGCTGTTTAT 7260
 Qy 7261 CCTAATGTTATCCCAATGTCAGCAAGAACTTGTGTGCTAGTAAAGAGAGGCTTCC 7320
 Db 7261 CCTAATGTTATCCCAATGTCAGCAAGAACTTGTGTGCTAGTAAAGAGAGGCTTCC 7320
 Qy 7321 AGATCAACAAAGACTGATTTCAATCTGATTTGAGGAGCCCACTTTTATAGTGA 7380
 Db 7321 AGATCAACAAAGACTGATTTCAATCTGATTTGAGGAGCCCACTTTTATAGTGA 7380
 Qy 7381 CTATATCTCTGAGTCTGATCTCTCTTTTAAATGAGAGACATTAATCCACATGG 7440
 Db 7381 CTATATCTCTGAGTCTGATCTCTCTTTTAAATGAGAGACATTAATCCACATGG 7440
 Qy 7441 CAGGCTGTGGGAGATTCAGAGATCAAAACGCTGTGATCAATCTGCTTCTGTTCC 7500
 Db 7441 CAGGCTGTGGGAGATTCAGAGATCAAAACGCTGTGATCAATCTGCTTCTGTTCC 7500
 Qy 7501 AGGCTCATCAGACTGGGCTTCTGAGCATGATTAACCATCCAGTCTTGGTACAGAA 7560
 Db 7501 AGGCTCATCAGACTGGGCTTCTGAGCATGATTAACCATCCAGTCTTGGTACAGAA 7560
 Qy 7561 CTGACACCAATCAACGAGACGCTGAGAGACTCTTGTCTTCAAGACAGACCTTGAAG 7620
 Db 7561 CTGACACCAATCAACGAGACGCTGAGAGACTCTTGTCTTCAAGACAGACCTTGAAG 7620
 Qy 7621 GGGCTGATGATCTGTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7680
 Db 7621 GGGCTGATGATCTGTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7680
 Qy 7681 CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCATCTAATCTTCAACTGTGCGGCT 7740
 Db 7681 CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCATCTAATCTTCAACTGTGCGGCT 7740
 Qy 7741 GACTTCTCTTCTTAACTGCGCAGATTAATGTTTGGCTTAACTGCGCTCATTAATCCG 7800
 Db 7741 GACTTCTCTTCTTAACTGCGCAGATTAATGTTTGGCTTAACTGCGCTCATTAATCCG 7800
 Qy 7801 CATCCCATCTCCAAATCTCACTGCTGATGATGATGATGATGATGATGATGATGAT 7860
 Db 7801 CATCCCATCTCCAAATCTCACTGCTGATGATGATGATGATGATGATGATGATGAT 7860
 Qy 7861 ATGCTGAGCGCATCAGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7920
 Db 7861 ATGCTGAGCGCATCAGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7920
 Qy 7921 TGCGCGCGCCAGATACCTGTATCGGTATGATGATGATGATGATGATGATGATGAT 7980
 Db 7921 TGCGCGCGCCAGATACCTGTATCGGTATGATGATGATGATGATGATGATGATGAT 7980
 Qy 7981 CTGCGAGATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 8040
 Db 7981 CTGCGAGATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 8040
 Qy 8041 TGGTGTAAAGCTCAGATTTTCAATCAATGCGGCTGCTGCTTATGATGATGATGAT 8100
 Db 8041 TGGTGTAAAGCTCAGATTTTCAATCAATGCGGCTGCTGCTTATGATGATGATGAT 8100
 Qy 8101 TGTGGTGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8160
 Db 8101 TGTGGTGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8160
 Qy 8161 ACCAGCTGTACGAGCAATCTCTCTCAAGTGTGATGATGATGATGATGATGATGAT 8220
 Db 8161 ACCAGCTGTACGAGCAATCTCTCTCAAGTGTGATGATGATGATGATGATGATGAT 8220
 Qy 8221 TTTGGCATTCAGTGGGCTGTTTTCAGAGATCCACTGATTTGAAAGCTTATTTGT 8280
 Db 8221 TTTGGCATTCAGTGGGCTGTTTTCAGAGATCCACTGATTTGAAAGCTTATTTGT 8280

QY 8281 CATGTCATCTAGTATGTTCCATTTTCCCTGTCGCGCTCTTAACAGAGTCGCAACCCCATCATTT 8340
 Db 8281 CATGTCATCTAGTATGTTCCATTTTCCCTGTCGCGCTCTTAACAGAGTCGCAACCCCATCATTT 8340
 QY 8341 TACTTCTTGTGGGCTCCTTTAGGAGCGCTCAAAATAGGACAACTGAAGCTGATTC 8400
 Db 8341 TACTTCTTGTGGGCTCCTTTAGGAGCGCTCAAAATAGGACAACTGAAGCTGATTC 8400
 QY 8401 CAGAGGCTCTGACAGACCGCTGAGTGATGAAAGTGAGAGGCTTCTCTACGAA 8460
 Db 8401 CAGAGGCTCTGACAGACCGCTGAGTGATGAAAGTGAGAGGCTTCTCTACGAA 8460
 QY 8461 ACCCTGAGAGCTCTGAGAGAGATTTGAGAGAGAGAACTCTGCGCTGTACAG 8520
 Db 8461 ACCCTGAGAGCTCTGAGAGAGATTTGAGAGAGAGAACTCTGCGCTGTACAG 8520
 QY 8521 AGGACTTTGAGAGCAATGCTGCTGCTGCAACCTTGACAAATTAATGATTTTCTTACG 8580
 Db 8521 AGGACTTTGAGAGCAATGCTGCTGCTGCAACCTTGACAAATTAATGATTTTCTTACG 8580
 QY 8581 TTCTGCTTCAAGAAATGCTCTCAGAGGTCCTCAAGGCTTACCA 8622
 Db 8581 TTCTGCTTCAAGAAATGCTCTCAGAGGTCCTCAAGGCTTACCA 8622

RESULT 2
 AB088140/c
 ID AB088140 standard; cDNA; 66494 BP.
 XX AB088140;
 AC
 XX 18-SEP-2002 (first entry)
 DT
 XX Human osteoblast differentiation related cDNA SEQ ID NO 47.
 DE
 XX Human; osteoblast; stem cell differentiation; bone tissue deposition;
 KW osteoporosis; osteopathic; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200250301-A2.
 PD
 XX 27-JUN-2002.
 PF
 XX 18-DEC-2001; 2001WO-US048276.
 XX
 PR 18-DEC-2000; 2000US-0255882P.
 PR 24-APR-2001; 2001US-0285691P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA (PROC) PROCTER & GAMBLE CO.
 XX
 PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
 PI Mertz L;
 DR WPI; 2002-557663/59.
 XX
 PT Use of genes and their expression profiles associated with osteoblast
 PT differentiation for screening modulators bone formation, for diagnosing
 PT or treating e.g. osteoporosis, or as markers for the differentiation
 PT process.
 XX
 PS Claim 1; SEQ ID NO 47; 78bp + Sequence Listing; English.
 XX
 CC The invention relates to genes and their expression profiles are used
 CC for: (a) screening modulators of precursor stem cell differentiation into
 CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
 CC deposition of bone tissue, abnormal rate of osteoblast formation or
 CC osteoporosis; or (c) treating or monitoring treatment of the conditions
 CC cited in (b), or monitoring the progression of bone tissue deposition.
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
 CC induced abnormalities in bone formation or bone loss, conditions that

CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
 CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
 CC or fibrous dysplasia. The present sequence is that of an osteoblast
 CC differentiation associated cDNA marker of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIDO
 CC at ftp.wido.int/pub/published_pct_sequences
 XX
 SO Sequence 66494 BP; 17680 A; 13908 C; 14466 G; 20440 T; 0 U; 0 Other;
 Query Match 16.2%; Score 1401; DB 6; Length 66494;
 Best Local Similarity 77.7%; Pred. No. 0;
 Matches 2034; Conservative 0; Mismatches 375; Indels 208; Gaps 20;
 QY 15 GTACCTTTAATCCAAAACGAGAAAGATAT-----ACAAAAGAAAATTAAGACC 68
 Db 64729 GTACCTTTAATCCAAAACGAGAAAGATATTAACAAACAAAAGAAAATTAAGACC 64670
 QY 69 AGTACCACTGATGAATATATCATGCAAGAAATCCCAACAAAATCTAGCTAACCCATCCA 128
 Db 64669 AATATCCCTAATGAATAATGATGCAAAAATCTCAGCAAAAATTTAGTGAACCAATCCA 64610
 QY 129 ACGACATATCAAGAAAGATATATCCACCATTTGCAAGTGGGTTTCATCAAGGGTG-CAAG 187
 Db 64609 ACAATATATCGAAAAGATATATCCACCATTTGCAAGTGGGTTTCATCAAGGGATGCCAGA 64550
 QY 188 ATAGGTTAATATCAATCAAGTCAATTAATGATATATCATCAATTAACAAATTAACAA 247
 Db 64549 ATAGGTTAATATCAATCAAGTCAATTAATGATATATCATCAATTAACAAATTAACAA 64490
 QY 248 AATACATGATCATCTCAATGATGCTGAAAAGCATTTGACAAATCTAATCTTCTT 307
 Db 64489 AATACATGATCATCTAATGATGCTGAAAAGCATTTGACAAATCTAATCTTCTT 64430
 QY 308 TATGATTAATACCTTACGAAAATGACATAGCAAGAAAGCATCTTAATGTAATTAAGC 367
 Db 64429 TATGATTAATACCTTACGAAAATGACATAGCAAGAAAGCATCTTAATGTAATTAAGC 64370
 QY 368 CATATATGACGAGCCACGACGAAACATTAATGATGAGGAGAAAGTGAAGAAACATGTC 427
 Db 64369 CATATATGACGAGCCACGACGAAACATTAATGATGAGGAGAAAGTGAAGAAACATGTC 64311
 QY 428 CCTGAACTGGAACAGACAGAGATG-CTACTTTCACCATCTTATTCATATAGTACT 486
 Db 64310 TATTAAGATGGAACAGAGAGAAAGATGCTTTCACCATCTTATTCATATAGTACT 64251
 QY 487 GGAAGTTTAAAGCCAGAGCATATGACAGAGAAAGAAATCAAGGCGCAACCAATTAATA 546
 Db 64250 GGAAGTTTAAAGCCAGAGCATATGACAGAGAAAGAAATCAAGGCGCAACCAATTAATA 64191
 QY 547 AGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATGATGATGATGATGATGATGATGAT 606
 Db 64190 AGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATGATGATGATGATGATGATGATGAT 64131
 QY 607 AGACTCATCCAGAAAGCTCTAGAACTGATATCATTAATTAATTAAGTTTCAGATACAA 666
 Db 64130 AGCTCATCCAGAAAGCTCTAGATCTGATGATGATGATGATGATGATGATGATGATGAT 64071
 QY 667 ACTAAATGACAAATCATGACCTGCTATATACCAACAGTGCACAACTGAGATCA 726
 Db 64070 AATCATGATATCAAAATCATGACCTGCTGCTTACCAACAGCAATCAAGTTGAGAAATCA 64011
 QY 727 AATCAAGAACTCAAAACATTTTACAAATAGCTGTAA-----AAAT 767
 Db 64010 AATCAATTAATCAAAACCTTTTACAAATAGCTGTAA-----AAAT 63951
 QY 768 ACTTAAGATATTTCTTACCAAGAGGTGAAGAGCTCTACAAAGAAATCTAACAACAC 827
 Db 63950 TCTTGGATATATCTTAACCAAGAGGTGAAGAGCTCTAACAAGAAATCTAACAACAC 63891
 QY 828 AGCT---GACATCATGATGACACAAACAGTGAACATCCCATGCTCATGATGG 883
 Db 63890 TGCTGAAGAAATCAAGGACACACAAATTAATTAAGAAACATCCCATGCTCATGATGG 63831

```

QY 884 GTAGATCAATATTGTGAAAAATGACCAATTTGCCAAAAGCAATCTACAGTTCAATGCAA 943
DB 63830 GTAGATCAATATTGTGAAAAATGACCAATTTGCCAAAAGCAATCTACAGTTCAATGCAA 63771
QY 944 TTCCCAACCAAAATATCTCATCTCTTTCACAGAACTTGAAAAAACAATTTCTAAATTC 1003
DB 63770 TGCCCAACCAAAATATCTCATCTCTTTCACAGAACTTGAAAAAACAATTTCCCAAGTC 63711
QY 1004 AT-----ATGAAACAAACCAACCAAAAAAATTTTCCGATGCCAAA 1050
DB 63710 ATGAAATTTTAGAACCAATGCTTAATCTGAAACCAAAAAAGAGCCCAATTTGTCAA 63651
QY 1051 GCAAGACTTAGCAAAAAGAAACAAATCTGAGGCACTCATTTCCCATCTTCAACTATAC 1110
DB 63650 GCAAG-CTAAGCAAAAAGAAACAAATCTGAG-ACACATTTGCTGACTTCAAACTATAC 63595
QY 1111 TACAGGCTTATATCACAACCAATCATGGAATGACATATAAAGTACGACATAGACCA 1170
DB 63594 CA-TAGGCAATAGTCAATCAAAAACATGATGCTGATTAAC-AGGCAATAGACCA 63537
QY 1171 TGGAAAAAGAGAGAAATCCAGAAATTAAGCAATATATTATAGCCAACTGATTTTGAC 1230
DB 63536 TGGAAAGAAATAGAGAACCAAAAATTAAGCAACACTTACAGCCCAATGATCTTCAAC 63477
QY 1231 AAAGCAACAAAAACATTAAGTGGGAAAGACATTTAGTTAACAAATGCTGAGAT 1290
DB 63476 AAAGCAACAAAAACATTAAGTGGGAAAGACCACTATTTCAACATGCTGCTGAGAT 63417
QY 1291 TATTGGCAACCAATGTCGAAAGAAATGAACTGATCTCTTGTCTTCACTTAATACAA 1350
DB 63416 AATTGGCAACCTATCATGTAGAAAGAAATTAACCTGAT-CTTCATCTTCAACGATTAATAA 63358
QY 1351 AATTGATACAAAGATGATCAAAAGACTTAATCTGAGACTTAAACCTTAATAATTTCTGA 1410
DB 63357 AATCACTCAAGATGATCAAAAGACTTAAACTTAAGCTGAACCAAGAAATTTCTAGA 63298
QY 1411 AGATTAACATGAAAAATGCTTCTAGACATTTCACTTGAAGCAAGCTTCATGCCCAAGAA 1470
DB 63297 AGATTAACATGAAAAATGCTTCTAGACATTTGCTTGAAGCAAGCTTCATGCAAGAA 63238
QY 1471 CCGAAAAGTAAATGCAAAAAACAAAAATTAATAGATAGACTTAATTAATCTTAATAAG 1530
DB 63237 CCGAAAAGCAATGCAAAAAACAAAAATTAATAGATAGACTTAATTAATACTTAATAAG 63178
QY 1531 CTTTGGCAGCAAAAACCAATCATTTAGCAGCAACCAAGCAACCCAGAGTAGAGAA 1590
DB 63177 C-TCTGACAGCAAAAACCAATCATTTAGCAGGTTAAACAGACAACTTATAGAG----- 63127
QY 1591 AATCTTCAAAAACCTAGCATCTGACTAAGACCTAATATCCGAAATCCAGAGAACTCAA 1650
DB 63126 -----TTTGGCAAACTTTCATCTGACAAAGGACTTAATCTCAGAACTTCAAAAGAAATCAA 63071
QY 1651 ACAAAATGAGCAAGAAAGAAACCAATCTCCATGAAAGAGTGGGCTTAAGGACATGAATAG 1710
DB 63070 ACAAAATGAGCAAGAA-----AAAAAATCCCATCAAAAAGTGGGCTTAAGGACATGAATAG 63016
QY 1711 ACAATTTCAAAAAGAAATATATACAAATGAGCAACAAAC--AGAAAAAATCTTAAATC 1768
DB 63015 ACAATTTCTCAAAAAGAAATATATACAAATGAGCAATTAACATATGAAAAAATCTCAACTT 62956
QY 1769 ACTAATGATTAAGGAAATGTAATCAACATGTAATGAGATACCACTTACTCTGCAAG 1828
DB 62955 ACTAGTATCAAGAAATGCAAAATCAAAACCAACATGCAATACCACTTACTCTGCAAG 62896
QY 1829 AATGCTATTAATTTTAAATTTTAAATTAATAGATGTTGGTCTGTGTGTGATTAAGG 1888
DB 62895 AATGCGATTAATTTTCAAAATTAATTAATGATGTTGGTGTGTGTGATTAAGG 62836
QY 1889 AACACTTTTACACTGCTGTGTGGAATGTAATCTTGCGCAACCACTATGAAAAACAGTGTG 1948
DB 62835 GACACTTTTACACTGCTGTGTGGAATTAATTAATCAACCACTACGGAACCAAGTGTG 62776

```

```

QY 1949 GAAATTTCTTAAGAACTTAATAAGTATGACCACTTTGATCCAGCAATCCCA----- 2000
DB 62775 GAGATTCCTTTAAGAACTTAATAAGTATGATCTACACTTTGACTGAGCAATCCCACTGCTGG 62716
QY 2001 ----- 2000
DB 62715 TATCTACCAAGCAAGCAAGATGATTAACAAAAAGATATCAATCAATGATGATTAAGAA 62656
QY 2001 ----- 2000
DB 62655 AATGATATTTTATATATATATAGTATACAGTATATACATATATATATATATATATA 62596
QY 2001 -----TTAATATGATTAATATATATATATATATATATATATATATATATATATATA 2057
DB 62595 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 62536
QY 2058 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2117
DB 62535 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 62476
QY 2118 GGGGTAATCAGAAATGAAAAACCAACATCATATGTTCTCACTTCAAGTGGGAGCTAA 2177
DB 62475 GAGTATCTCAGGAATGAAAAACCAACATCATATGTTCTCACTTCAAGTGGAGCTAA 62416
QY 2178 GCTGT---GAGGACAGAAAGCAT-AGATGATTAATGAATCTGGGAGCTTGAAGGGA 2233
DB 62415 GCTGTGAGGAGGAGCGCAAGGACATTAAGCAATTAATGAATCTTGGGAGCTTGAAGGGA 62356
QY 2234 AGATGGAAGAGAGGCGAGGAGTAAAGACTACACATGAGGTACAGTGTACATGCTCAG 2293
DB 62355 AGGATGGAAGAGTCAAGGAGTAAATTAAGACTACACATGAGGTAGAGTACATGCTCAG 62296
QY 2294 GTGATGGTCCACCAAAATCTCAGAAATTAACAATAAGAACTTATCATGAGAAAC 2353
DB 62295 CTGAGAGGTCAACCAACATCTCAGAAATGCGCCCTAATAACTTATCATGATTAACAAAC 62236
QY 2354 ACCACTGTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTTTC 2413
DB 62235 ACCACTGTTGCCCCCAAAAGCATGTAAT-AAAATATATATATTAATAATTAATTAATTA 62177
QY 2414 ACAGATTTAAAAAGTTCACTGTTGAGAGTTTAA 2450
DB 62176 ATTGCATGAGGTAGGTTTATTTATCTCTGTTTACAA 62140

RESULT 3
ABA90193 0/C
WP Sequence split into 5 fragments LOCUS ABA90193 Accession ABA90193
WP Fragment Name Begin End
ABA90193_0 1 110000
ABA90193_1 100001 210000
ABA90193_2 200001 310000
ABA90193_3 300001 410000
ABA90193_4 400001 465237
ID ABA90193 standard; DNA; 465237 BP.
XX
AC ABA90193;
XX
DT 11-FEB-2002 (first entry)
XX
DB Human oestrogen receptor alpha gene.
XX
KW Human; oestrogen receptor alpha; ESR-alpha; ER; chromosome 6; Syme-2;
KW synaptic nuclei expressed gene 2; haploids; cytostatic; osteopathic;
KW cardiac; vasotropic; gene therapy; vaccine; cancer; osteoporosis;
KW cardiovascular disease; oestrogen receptor; de.
XX
OS Homo sapiens.
XX
MOZ00162969-A2.
XX
PD 30-AUG-2001.
XX

```


FT	/*tag= b	
FT	/number= 5	
FT	replace(168054,G)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	168121..169542	
FT	/*tag= b	
FT	/number= 5	
FT	/cons_splice= (5'site:YES,3'site:NO)	
FT	169543..169825	
FT	/*tag= b	
FT	/number= 6	
FT	replace(169812,G)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	replace(169823,G)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	169826..169866	
FT	/*tag= b	
FT	/number= 6	
FT	/cons_splice= (5'site:YES,3'site:NO)	
FT	169867..170678	
FT	/*tag= b	
FT	/number= 7	
FT	replace(170035,A)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	replace(170256,C)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	replace(170358,G)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	replace(170487,C)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	170679..204911	
FT	/*tag= b	
FT	/number= 7	
FT	204912..205102	
FT	/*tag= b	
FT	/number= 8	
FT	205103..242969	
FT	/*tag= b	
FT	/number= 8	
FT	242970..243086	
FT	/*tag= b	
FT	/number= 9	
FT	replace(243055,T)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	243087..306167	
FT	/*tag= b	
FT	/number= 9	
FT	replace(243187,C)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	306168..306503	
FT	/*tag= b	
FT	/number= 10	
FT	replace(306292,A)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	replace(306382,G)	
FT	/*tag= a	
FT	/standard name= "Single nucleotide polymorphism"	
FT	306504..373639	
FT	/*tag= b	
FT	/number= 10	

FT	exon	373640..373778
FT	/+tag= b	
FT	/number= 11	
FT	373779..422963	
FT	/+tag= b	
FT	/number= 11	
FT	422964..423097	
FT	/+tag= b	
FT	/number= 12	
FT	replacement(423067,C)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	423098..456353	
FT	/+tag= b	
FT	/number= 12	
FT	replacement(423149,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(423163,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(423220,A)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(423232,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(423258,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	456354..456537	
FT	/+tag= b	
FT	/number= 13	
FT	456538..460700	
FT	/+tag= b	
FT	/number= 13	
FT	replacement(459706,C)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(459832,A)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(459913,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(460024,G)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(460056,T)	
FT	/+tag= a	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	replacement(460159,G)	
FT	/+tag= a	
Query Match	16.1%; Score 1390.4; DB 6; Length 110000;	
Best Local Similarity	77.8%; Pred. No. 1.6e-312;	
Matches 1970; Conservative	0; Mismatches 381; Indels 181; Gaps 17;	
Dt	1 TGTATGAAGCCAAATGCTACTTTAATACCAAAACAGAAAAGATATACAAAAAGAAAC	60
Dt	7473 TCTATGAAGTACAGATTACTTAATATACCMAAACAGAAAGACAT--AACCGAAAAAC	7417
Oy	61 TATAGACCAATACCACTGATGAATATATACATGCAGAAAAATCCCAACAATAATCTAGTAAAC	120
Dt	7416 TACAGACCAATATCCCTGACGAACATGATCTTAAATCCTTAACAAAATATCTAGCTAAC	7357
Oy	121 CCAATCCAACAGCATATACAGAAGATTAATCCACATTTGTCAGTGGGGTTTCATACCAAGG	180
Dt	7356 TGAATCCAACAACATATCAAAAAAGTATATTCACCATGATCAAGTGGGTTTCATACCAAGG	7297
Oy	181 GTGCA-GGATAGGTTTAACATACACAGTCAATTAATGTGATATCATCATATAACGAATT	239
Dt	7296 ATGCAGGGATGGTTTAACATACACAGTCAATTAATGTGATATCATCATATAACGAATT	7237

QY 240 AAAAAAATCAATGATCATCTCATAGTGTGAAAAAGCAATTTGACAAAATCTAA 299
 DB 7236 AAAAAAATCATATATCATCTCAACAGATGAGAAAAAGCAATTTGACAAAATCTAG 71177
 QY 300 CATTTCTTATGATTTAAACCTTGAGCAAAAATCGACATAGAAAGACATACCTTATGTA 359
 DB 7176 TATCCCTTATGATTTAAACCTTGAGCAAAAATCATCTTCAAGGACATCTCAATTA 71117
 QY 360 ATAAAAGCATATATGACGAGCCACAGCAAAATTATTAATGATGGGAAAAAGTTGAA 419
 DB 7116 ATAAAAGCATATATGACAAACCCACAGCAAACTAATATGATGGGAAAAAGTTGAA 7057
 QY 420 ACATTTCCCTGAGAACTGAGAAAGCAAGAGAT-GCTACTTTCACCACTTCTATCAAC 478
 DB 7056 GGAATTCCTTGAGAACTGAGAAAGCAAGAGATGGCAATTCACCACTTCTTCAAC 6997
 QY 479 ATAGTAGGAAAGTTTATGAGCAAGCAATCAGACAGAAAGAAATCAAGGACCCCA 538
 DB 6996 ATAGTAGGAAAGTTTATGAGCAAGCAAGCAAGAAAGAAATTAAGGACATCCA 6937
 QY 539 ATCAATTAAGGAAAGTCAAACTGTCTCTTCACTGATGATGATTTATACCTAGAA 598
 DB 6936 ATTGACGAAGAGAAAGTGAATTTGTCTGCTGCAATATGATTTTACTTCAA 6877
 QY 599 AACCTTAAGACTATCCAGAAAGCTCTTAAATCTGATACATTAATTGATGAAATTTCA 658
 DB 6876 AACCTTAAGACTCTCCAGAAAGCTCTTAAATCTGATGAAAGAAATTTGAGAAATTTTC- 6818
 QY 659 GGATACAAACTTAATGTGACAAATCAGTAGCACTGTATACCAAGTAGCAAGCT 718
 DB 6817 GGATACAAAGTAAATGTGACAAATCAGTAGCTTCTTATACCAAGTAGCAAGCT 6758
 QY 719 GAGATCAAAATCAAGAACTCAAACTTTTACAATAGCTGTAA-----AAATATC 769
 DB 6757 GAGATCAAAATCAAGAACTCAAACTTTTATGATAGCTGCAAAAGATGAAATTAATATGC 6698
 QY 770 TTAAGATATTTCTTAACCAAGAGAGTG-AAGAACTCTACAGAGAAATCTCAAAACCA 828
 DB 6697 TTAAGATATTTCTTAACCAAGAGAGTGAAAGAAATCTGACAGAGAAATCTCAAAACCT 6638
 QY 829 GCTGACATCATAGATGACAAACAAAGTAGAAACATCCCATGCTCATGGAGGTGA 888
 DB 6637 GCTGAAAGAAATCATAGACGACAAATGAAACATCTCCATGCTCATGATGGGTGA 6578
 QY 889 ATCAATATTTGAAATGACCAATATTTGCAAAAGCAATCTAACATTCATATCCC 948
 DB 6577 TTCAATATTTGAAATGAC-----TGCAAAAGCAATCTCAATTCATGCAATCCCC 6523
 QY 949 ACCAAATATCATCATCTTTCTTCAAGAACTAGAAAAAAACATTTCTAAATCATATG 1008
 DB 6522 ATTAAAT--ACCAACATTTCTTCAAGAACTTG-AAGACAAATTTCTAAATCATATGCG 6467
 QY 1009 GAACAAACACCAAAAAAAGAAAAAAGCCGCAATGCAAGCAAGCAAGCTTAGCAAAAG 1068
 DB 6466 GAAC-----CBAHAHAAGCTGCTATAGCAAGCAAGCAAGCTTAGCAAAAG 6421
 QY 1069 AACAAATCTGAGGACATCATTTTCCATCTTCAATATATCTCAAGGCTATATATCC 1128
 DB 6420 AACAAATCTGAGGACATCATCTTCAAGATTTCAACATGATTTAAGGCCACGTACCC 6361
 QY 1129 AAAACATCATGACCTGACATTAATGAGCAATGAGCAATGAGCAAGAAAGAGAGAT 1188
 DB 6360 AAAACATCATGATCTGTATTAATTAAGGCAATGAGCAATGAGCAAGATGATGATAC 6301
 QY 1189 CCAAAATTAAGCCAAATATTTATAGCAATGATTTTGAAGCAAGCAAGCAAGCAAGCTA 1248
 DB 6300 CCAAAATTAAGCCAAATATTTATAGCTATGATCTTCAAGCAAGCAAGCAAGCAAGCTA 6241
 QY 1249 AAGTGGGAAAAAGCATTTAGTTTACAAATGATGATGATTTATTTGCAAGCCACATGT 1308
 DB 6240 AAGTGGGAAAAAGCACCTTTTGAACAAATGATGATGATTTATTTGCTAGCCGAATGT 6181

QY 1309 GGAAGATGAATGAACTGGAATCCCTTGCTCTCACTTAATCAAAAAATGATAGAGATGAT 1368
 DB 6180 AGGAAATGAATGAACTGGAAT-TCTCAATTTCTCACTTAATCAAAAAATCACTTAGATGAT 6122
 QY 1369 CAAGACTTAATGATGAGACCTTAAGAAATTTCTAGAAAGTAAATCATCAGAAAT 1428
 DB 6121 TAAGACTT-----TCAGACTGGAATCTTAAGAAATTTCTAGAGATTAACATCGAAAAAC 6067
 QY 1429 GCTTCTGACATTCATCTTAGGCAAGACTTCATGAGCAAGAAACCAAAAGTAAATGCAAC 1488
 DB 6066 CTTTCTGACATTCATCTTAGGCAAGACTTCATGAGCAAGAAACCAAAAGTAAATGCAAT 6007
 QY 1489 AAAAAAATTAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1548
 DB 6006 AAAAAAATTAATTAATGAGCTGGAATCTTAATTAATTAATTAATTAATTAATTAATTAAT 5947
 QY 1549 AATCATTAAGCAGACCAACAGACCAAGCAAGCTGAGAGAAATCTTCAAACTAAGC 1608
 DB 5946 AATCATTAAGCAGACCAACAGACCAAGCAAGCTGAGAGAAATCTTCAAACTAATAC 5887
 QY 1609 ATCTGACTAAGACTTAATATCTGGAATCTCAGAGAACTCAAGAAATCAGCAAGAAAGAA 1668
 DB 5886 ATCTGACTAAGACTTAATATCTGGAATCTCAGAAATCTCAAACTAAGCAAGCAAGCAAG 5832
 QY 1669 AGCAAAATCCCATGGAAGAGTGGCTTAAGAGACATGAATGACAAATTTCTCAAAAGAA 1728
 DB 5831 TTTAAATATCTCTTCAAGAAAGTGGCTTAAGAGACATGAATGACAAATTTCTCAAAAGAA 5772
 QY 1729 TATCAAAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1784
 DB 5771 TATCAAAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5712
 QY 1785 ATGTAATCAACACTGTAATGAGTACCACTTACCTCTGCAAGAAATGATTAATTTAA 1844
 DB 5711 ATGTAATCAACACTGTAATGAGTACCACTTACCTCTGTAAGAAAGCAAGCAAGCAAGCAAG 5652
 QY 1845 AATCTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 DB 5651 AATCTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5592
 QY 1905 TGTGGAATGTAATCTTGGCAACCACTATGAGAAACAGTGTGAAATTTCTTAAGGAA 1964
 DB 5591 TGTGGAATGTAATCTTGGCAACCACTATGAGAAACAGTGTGAAATTTCTTAAGGAA 5532
 QY 1965 CTAAGATGATGAGCAATTTGATTCAGCAATCCAT-----A 2003
 DB 5531 CTAAGATGATGAGCAATTTGATTCAGCAATCCATCTTCTAGAGGAA 5472
 QY 2004 AATATGATTAATTAATTA----- 2021
 DB 5471 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5412
 QY 2022 ----- 2021
 DB 5411 ATTGCAAAATCTGTGAAACCAACCAATGCCATCAATCAATGATGATTAAGATTTCTG 5352
 QY 2022 ---TATTTATATACATGAGATTAACATCACTCAGCCATTAAGAAATTAATGATGATTC 2078
 DB 5351 TAGTATATATATATATAGGATTAATCACTCAGCTATTAAGAAATTAAGGCAATTC 5292
 QY 2079 ACAGCAATCTAGATGAAATTTGAGACCTTATTTTAAGTGGGTTAACTCAGGAATGGA 2138
 DB 5291 GCAATGACCTGAGAGATGAGATTTGAGACCTTATTTTAAGTGAATTAATCAGGAATGGA 5232
 QY 2139 ACCAAATCATATATGATCTCACTTACAGAGTGGGCTTAAGTGTGAGAGCAAGAGCAT 2198
 DB 5231 AACAAATCATATATGATCTCTATGATTTATAGGAGCTTAAGAGATGAGCAAGAGCAT 5172
 QY 2199 AGAATGATTAATTAATGATCTTGGGAGCTTGAAGGAGAGATGAGAGAGAGGAGGATTA 2258
 DB 5171 AATATGATTAATTAATGATCTTGGGAGCTTGGAGGAGAGGATTTGAG-GGGCAAGAGATTA 5113
 QY 2259 AAGACTACAAATGAGGATGAGTACCTGCTCAGGTGATGAGGTGACCAAAATCTCAGA 2318

FT	/tag= s
FT	/number= 4
FT	replace(306292,A)
FT	/tag= at
FT	/standard_name= "SNP"
FT	replace(306382,G)
FT	/tag= au
FT	/standard_name= "SNP"
FT	306504. .373639
FT	/tag= t
FT	/number= 4
FT	373640. .373778
FT	/tag= u
FT	/number= 5
FT	373779. .422963
FT	/tag= v
FT	/number= 5
FT	422964. .423097
FT	/tag= w
FT	/number= 6
FT	replace(423067,C)
FT	/tag= av
FT	/standard_name= "SNP"
FT	423098. .456353
FT	/tag= x
FT	/number= 6
FT	replace(423149,G)
FT	/tag= aw
FT	/standard_name= "SNP"
FT	replace(423163,G)
FT	/tag= ax
FT	/standard_name= "SNP"
FT	replace(423220,A)
FT	/tag= ay
FT	/standard_name= "SNP"
FT	replace(423232,G)
FT	/tag= az
FT	/standard_name= "SNP"
FT	replace(423258,G)
FT	/tag= ba
FT	/standard_name= "SNP"
FT	456354. .456537
FT	/tag= y
FT	/number= 7
FT	456538. .460700
FT	/tag= z
FT	/number= 7
FT	replace(459706,C)
FT	/tag= bb
FT	/standard_name= "SNP"
FT	replace(459832,A)
FT	/tag= bc
FT	/standard_name= "SNP"
FT	replace(459913,G)
FT	/tag= bd
FT	/standard_name= "SNP"
FT	replace(460024,G)
FT	/tag= be
FT	/standard_name= "SNP"
FT	replace(460056,T)
FT	/tag= bf
FT	/standard_name= "SNP"
FT	replace(460159,G)
FT	/tag= bg
FT	/standard_name= "SNP"
FT	replace(460553,C)
FT	/tag= bh
FT	/standard_name= "SNP"

Query Match 16.1%; Score 1390.4; DB: 8; Length 110000;
 Match Local Similarity 77.8%; Pred. No. 1.6e-312;
 Matches 1970; Conservative 0; Mismatches 381; Indels 181; Gaps 17;

QY	1	IGTATGAGCCAAATGTCCTTTAATACCAAAACCGAAGAAAGATTTACAAAAAGAAAC	60
Db	7473	TCATGAAGTCAAGCATTACTTTAATACCAAAACCGAAGAAAGACT--AACGAAAAAC	7411
QY	61	TATAGACAGTACGATGATGATTAATACATGACAGAAATCCGCCAACAAATATCTAGCTAAC	120
Db	7416	THGAGCCAAATATCCCTGACGACATGATGCTAATAATCTTTACAAATATCTAGCTAAC	735
QY	121	CCAAATCCAAAGCATATCAAGAAAGTAATCCACGATTTGTCAGTGGGTTTCATACGAGG	180
Db	7356	TGAATCCAAACAATATCAAAAAGTAATTCACCATGATCAAGTGGGTTTCATACGAGG	729
QY	181	GTGCA-GGATAGGTTTAACATACACAGTCATTAATGTGATACATACATATAACGAATT	239
Db	7296	ATGCAAGGATGGTTTAAATACATACAGAGTCATTAATGTGATACCATATAACGAATT	723
QY	240	AAAAACAAAATTCATGATCATCTCAATAGATGCTGAAAAGATTGACAAATCTAA	299
Db	7236	AAAAACAAAATTCATATATCATCTCAACAGATGACAAAAGCATTTGACAAATCTAG	717
QY	300	CATTTCCTTATGATTTAAACCTTGACGAAAATCGACATAGAAAGACATACCTTAATGTA	359
Db	7176	TATCCCTTTATGATTTAAACCTCTGACGAAATCATCATTCAGGAGCATTACTCAATATA	711
QY	360	ATPAAAGCCATATATGACGACCCACGACAAACATTATCTGAAATGGGAAAAGTTGAA	419
Db	7116	ATPAAAGCCATATATGACAAACCCACGACCAACATTATCTGAAATGGGAAAAGTTGAA	705
QY	420	ACATTGCCCTGAGAACTGGAACAAAGCAAGAT-GCTACTTTACCACTTCTATTCAAC	478
Db	7056	GGATTCCCTGAGAACTGGAACAAAGCAAGATGGCCATTCTCACCTCTCTTCAAC	699
QY	479	ATAGTAGTGAAGTTTATGACGAGACATCAGACAAAGAAAGAAATCAAGGGACCCA	538
Db	6996	ATAGTAGTGAAGAAATCTTAGCTAGAGCAACAGACAAAGAAAGAAATPAAAGGCATCCA	693
QY	539	ATCAATPAAAGGAAGTCAACTGTCCTGTTCACTGATGATGATTGTAATACCTAGAA	598
Db	6936	ATTGACGAAGGGAAGTGAAATTTGCGCTGTTCTGCAATATGATTTGTTACTTCAA	687
QY	599	AAACCTTAAGCTCATCCGAAACCTCCTTAACTGATACATAATTCAGTAAAGTTTCA	658
Db	6876	AAACCTTAAGCTCCTCCGAAACCTCCTTAACTGATPAAAGAAATTCAGGAAAGTTTC-	681
QY	659	GGATACAACTPAAATGTACCAAAATCTAGTACACTGCTATACCAACAGTGACCAAGCT	718
Db	6817	GGATACAAAGTAAATGTACCAAAATCTAGTACCTTTCTATACCAACAGTGACCAAGCT	675
QY	719	GAGATCAAAATCAAGAACTCAAAACCTTTTCAATAGCTGTAA-----AAATATC	769
Db	6757	GAGATCAAAATCAAGAACTCAAAACCTTTTATGATAGCTGCAAAAGATGAATATAATATGC	669
QY	770	TTAAGATATTTCTTAACCAAGAGAGTG-AAGAGCTCTACAAAGAAAACTACAAACACA	828
Db	6697	TTAGAAATATCTTAACCAAGAGAGTGAAAACTGTACAAAGAAAACTACAAACACT	663
QY	829	GCTGACATCATAGATGACACAAACAGTGGAAACACATCCCATGCTATGATGGGTGAG	888
Db	6637	GCTGAAAGAAATCATAGACACACAAATGGAACATCCCATCTCATGTGATGGGTGAG	657
QY	889	ATCATATTTGTGAAAATGACCATATTGCGAAAGCAATCTACAGTTCAATGACATTTCC	948
Db	6577	TTCAATATTTGTGAAAATGAC-----TGCAAAAGCAATCTACAAATTCATGACATCCC	652
QY	949	ACCAAAATATCATCATTTTTCACGAGACTAGAAAAAAACAATTTCAAAATTCATATG	1008
Db	6522	ATTAAAT--ACACACATTTCTACAGAACTTG-AAGACCAAAATTTCAAAATTCATATG	646
QY	1009	GAAACAACAACAAAAAAGAAAAAAGCCGATAGCGAAAGCAAGCTTAGCAAAAG	1068
Db	6466	GAAC-----CCAAAAAGCTGCAATAGCGAAAGCAAGCTTAGCAAAAG	642
QY	1069	AACAATCTGAGGATCATCAATTACCAATCTTCAACTATACAAAGCTATATATCAC	1128

```

Db 6420 AACAAATCTGGAGGCTACCTACCAAGTTTCAATCTATTTATTAAGGCCACAGCTACC 6361
Qy 1129 AAAACATCATGGCACTGACATTAATACTAGGCACTAGACCAATGAGAAAGAGAGAAAT 1188
Db 6360 AAAACAGCATGTGTACTGTATTAATAATAAGGACATAGACCAATGAGAAAGAGATTAAC 6301
Qy 1189 CCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAACCAAAACAAATTAATTA 1248
Db 6300 CCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAACCAAAACAAATTAATTA 6241
Qy 1249 AAGTGGGAGAAAGACATCTAGTTAACAATGCTGAGATTAATTTGGCAAGCCACATGT 1308
Db 6240 AAGTGGGAGAAAGACATCTTTTGAACAATGCTGAGATTAATTTGGCAAGCCACATGT 6181
Qy 1309 GGAGAAATGAACCTGATCCCTGTCTCTCACTTAATTAACAATAATGATTAACAAGATGAT 1368
Db 6180 AGGAGAAATGAACCTGAT -TCATATTTCTCACTTAATTAACAATAATGATGATGAT 6122
Qy 1369 CAAGACCTTAATCTGAGACTTAACCAATTAATTAATTAATTAATTAATTAATTAATTA 1428
Db 6121 TAAAGACTT-----TCAGACTGAACTCTAATAATTTAGAGATTAACACTGAAATTAAC 6067
Qy 1429 GCTTGAACATTCCTTAGGCAAGAACTTCAATGAGCAAGCAACCAAAAGTTAAATGCAAC 1488
Db 6066 CTTTCTAGCATTTGGCTTAGGCAAGAACTTCAATGAGCAAGCAACCAAAAGTTCAAT 6007
Qy 1489 AAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1548
Db 6006 AAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5947
Qy 1549 AATCATTAAGCAAGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1608
Db 5946 AACACTGACAGCAAGTAACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5887
Qy 1609 ATCTGACTAAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1668
Db 5886 ATCTGACTAAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5832
Qy 1669 AGCAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1728
Db 5831 TTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5772
Qy 1729 TATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1784
Db 5771 TATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5712
Qy 1785 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1844
Db 5711 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5652
Qy 1845 AATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1904
Db 5651 AATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5592
Qy 1905 TGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1964
Db 5591 TGTGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5532
Qy 1965 CTAAAGTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2003
Db 5531 CTAAAGTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5472
Qy 2004 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2021
Db 5471 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5412
Qy 2022 ----- 2021
Db 5411 AATTGCAAAATCTGGAACCAACCAATGCCATCAATGATGATTAATTAATTAATTAATTAATTA 5352
Qy 2022 ---TATTTATTAATCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2078
Db 2022 ---TATTTATTAATCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2018

```

```

Db 5351 TAGTATATATATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5292
Qy 2079 ACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2138
Db 5291 GCAGTACCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5232
Qy 2139 ACCAAATCATATGTTCTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2198
Db 5231 AACAAATCATATGTTCTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5172
Qy 2199 AGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2258
Db 5171 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5113
Qy 2259 AAGCTACCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2318
Db 5112 AATACGCAAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5053
Qy 2319 AATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2378
Db 5052 AATCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4993
Qy 2379 GAATTAATAATA 2390
Db 4992 GAAAAAGAAAA 4981

RESULT 6
ABX08336_13
Continuation (14 of 17) of ABX08336 from base 1300001 (Human phosphodiesterase 4D (PDE4D
WP Sequence split into 17 fragments LOCUS ABX08336 Accession ABX08336
WP Fragment Name Begin End
WP ABX08336_00 1 110000
WP ABX08336_01 100001 210000
WP ABX08336_02 200001 310000
WP ABX08336_03 300001 410000
WP ABX08336_04 400001 510000
WP ABX08336_05 500001 610000
WP ABX08336_06 600001 710000
WP ABX08336_07 700001 810000
WP ABX08336_08 800001 910000
WP ABX08336_09 900001 1010000
WP ABX08336_10 1000001 1110000
WP ABX08336_11 1100001 1210000
WP ABX08336_12 1200001 1310000
WP ABX08336_13 1300001 1410000
WP ABX08336_14 1400001 1510000
WP ABX08336_15 1500001 1610000
WP ABX08336_16 1600001 1691080

Query Match 16.0%; Score 1375.6; DB 6; Length 110000;
Best Local Similarity 77.1%; Pred. No. 4.7e-309; Indels 182; Gaps 19;
Matches 1993; Conservative 0; Mismatches 409;

Qy 1 TGTATGAGCAATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 58
Db 25493 TGTATGAGCAATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25552
Qy 59 ACTATGAGCAATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118
Db 25553 ACTATGAGCAATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25612
Qy 119 ACCAATTCACAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 178
Db 25613 ACTGAATTCACAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25672
Qy 179 GGGTGAG-GATAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 237
Db 25673 GGGTGAGAGATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25732
Qy 238 TTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 297
Db 25733 TTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25792

```

QY 298 AACATTTCTTATGATTTAAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 357
 Db 25793 AGCATCCCTTTATGATTTAAAGCTCTCAGCAAAATCAGCATACAAAGGACATACATTAATG 25852
 QY 358 TAAATAAAGCATATATGACGGACCCACAGCAAACTTATCTGAATGGGAAAAAGTTGA 417
 Db 25853 TAAATAAAGCATATATGACAAACCCACAGCAAGCTAATCTGAATGGGAAAAAGTTGA 25912
 QY 418 AAACATTTCTCCTGAGAACTGGAACAAAGCAAGATG-CTACTTTCAACCTTCTATTCA 476
 Db 25913 AAGAAATTCCTCTGAGAACTGGAACAAAGCAATGATGCTCCTCACCACCTTCTTCA 25972
 QY 477 ACATATGATGTAAGTTTAAAGCCAGAGCAATCAGCAAGAAAAAGAAATCAAGGCGACC 536
 Db 25973 ACATATGATGTAAGTTTAAAGCCAGAGCAATCAGCAAGAAAAAGAAATCAAGGCGACC 26032
 QY 537 AAATCAATTAAGAGAGTCAAACTGTCCCTGTTCACTGATGATGATTTGATTAACCTAG 596
 Db 26033 AAATCGTAAGAGAGTCAAACTGTCTCCTGTTGCTGATGATTAATTAATTAACCTAG 26092
 QY 597 AAAACCCCTAAGAGTCTATCCAGAAAGCTCCTAGAACTGATACATTAATTCAGTAAAGTT 656
 Db 26093 AAAACTCTAAGAGTCTCTCCAGAAAGCTCCTAGAACTGATTAATTAATTCAGAAAGTT 26152
 QY 657 CAGGATACAACTAATATGACAAATCAGTACGCTCTATACCAAGTGAACCAAG 716
 Db 26153 CTGATATCAAGATTAATGTAACAATCAGTACGCTCTCTATATACCAAGTGAACCAAG 26212
 QY 717 CTGATATCAAGATTAAGAACTCAAACTTTTCAATGCTGTAA-----AAA 765
 Db 26213 GGGGAATCAAAATCAAGAACTCAAACTTTTCAATGCTGTAA-----AAA 26272
 QY 766 ATACTTAAGATATTTCTTACCCAGAGAGTGAAGAGCTCTACAGAAAGAACTCAAAAC 825
 Db 26273 ATACTTAAGATATTTCTTACCAAGAGAGTGAAGAGCTCTACAGAAAGAACTCAAAAC 26332
 QY 826 ACAGCTGA-----CATCATGATGACACAAACAGTGAACACATCCCATGCTCAGAT 881
 Db 26333 ACTCTGAAGAGATTCATAGACAAATACAAACAAATGGGACACATCCCATGATCAGAT 26392
 QY 882 GGGTGAATCAATTTTGTGAAAAATGACATATTTGCCAAAGCAATCTTCAAGTTCAATGC 941
 Db 26393 GGGTGAATCAATTTTGTGAAAAATGACATATTTGCCAAAGCAATCTTCAAGTTCAATGC 26452
 QY 942 AATCCCAACCAAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACAATTTCAAAAT 1001
 Db 26453 AATCCCAACCAAAATATCAACATCATCTTCTTCAAGAAATAG-AAAAACAATTTCAAAAC 26511
 QY 1002 TCATATGGAACAAACCAAAAAAAGAAAAAAGCCGCAATGCAAGCAAGCACTTAG 1061
 Db 26512 TCATATGGAAC-----CAAAAAAGAGCTGCAATAGCAAAAGCAAGCAAG 26557
 QY 1062 CAAAAAAGCAAAATCTGAGAGCATCACTTACCATCTTCAACATATCTTCAAGGCTAT 1121
 Db 26558 CAAAAAAGCAAAATCTGAGAGCATCACTTACCATCTTCAAAATTAATCTTAAAGCCAT 26617
 QY 1122 AATCAACCAAAACATCTGAGCACTGACATTAAGCACTAGACCAATGAAAAAGAG 1181
 Db 26618 AGTCAACCAAAATGAGCATGCTGTATTAATAATGACATATGACCAATGAAACAGAT 26677
 QY 1182 AGAGATTCAGAAATTAAGCCAAATTAATTAATGCACTGATTTTGAACAAGCAAA 1241
 Db 26678 AGAGATTCAGAAATTAAGCCAAATTAATTAATGCACTGATTTTGAACAAGCAAA 26737
 QY 1242 AAAATTAAGTGGGAAAAA-CACTTCATGTTAACTAATGATGCTGAGATTAATTTGGCAG 1300
 Db 26738 AAAATTAAGTGGGAAAAAAGATTAACCTTTCAACAAATGATGCTGAGATTAATTTGGCAG 26797
 QY 1301 CCAGATGTAAGATGAATGAACTGATCCCTTGTCTCTCACTTAATCAAAATTAATGATCA 1360
 Db 26798 CCAGATGTAAGATGAATGAACTGATCC--TATCTCTCACTGTATCAAAATTAATCACTCA 26855

QY 1361 AGATGATCAAAAGCTTAAATCTGAGACTTAAACCATAAATTTCTAGAAATTAACATC 1420
 Db 26856 AGATGATTAAGGGCTTAAACCTAAGACGTCAAACTAT-GAAATTTTGAAGATTAACCTT 26914
 QY 1421 AGAAAAATGCTTCAAGCATCTTATGGGCAAAAGCTTCAAGGCAAGAACCCAAAGTA 1480
 Db 26915 GAAAAAACCTTCTTCAAGCATTTGCTTGAAGCAAGATTTCAAGACCAAGAACCCAAAGTA 26974
 QY 1481 AATGCAACAAAAAATAATTAATGATAGGACTTAAATTAAC-TAAAAAGCTTTTGGCC 1539
 Db 26975 AATGCAATTAATAAATAATTAATGATAGGACTTAAATTAACCTTTACAGACTTTTGGAG 27034
 QY 1540 AGCAAAAACATCAATTAGCAGCAAAACAGCAACCAACCGAGTGAGAG-AAATCTTCA 1598
 Db 27035 GGCAAAAAGCAACATGACAGAGTAAACAGCAACCAACAGAGTGGGAAAAATCTTCA 27094
 QY 1599 CAATCTAAGCATCTGACTAAGGACTAATATCCGGAATCCAGAAAGACTCAACAAATCA 1658
 Db 27095 CAATCTAAGCATCTGACTAAGGAG-GTAGTATCAAGATTTCAAGAGACCCCAACAAATCA 27153
 QY 1659 GCAAGAAAGAAACAAACATCCATGAAAGAGTGGCTAAGGACATGATAGCAATTC 1718
 Db 27154 GTAAAGAAAAAACAACATCCATCAAAAGTAAAGGCTAAGGACATGATGAGCAATTC 27213
 QY 1719 CAAAAAGATATCAAAATGGCCCAACAC--AGAAAAATGCTTAACTATCACTAATGA 1776
 Db 27214 CAAAAAGATATCAAAATGGCCCAACACATATGAAAAATGCTCAACATCACTAATGA 27273
 QY 1777 TTAGGAAATGTAATTAACAC-CTGTAATCCGATACACACTTCTCCTGCAAGAAATGT 1834
 Db 27274 TCAAGAAATGTAATTAACACAAATGTAATACACGTAATCTCTGCAAGAAATGT 27333
 QY 1835 CATATTTTAAATATCT--AAAAATTAATGATGTTGGTGTCTGTGATTAAGAAACA 1892
 Db 27334 CATATTTTAAATATTTTAAATTAATGATGTTGGTGTGATGAGAGGATGATCAGAAACA 27393
 QY 1893 CTTTCACTGCTGTGGGAAATGTAACCTTGCGCAACCATATGAAAAACAGTGTGAAA 1952
 Db 27394 CTTTCACTGCTGTGGGAAATGTAACCTATGTAACGACATTAATGAAAAAGTGTGGGA 27453
 QY 1953 TTTCTTAAAGAACTTAAAGTATGATGCAACCTTATATCAGAAATCCAT----- 2001
 Db 27454 TTCTTAAAGAACTTAAAGTATGATGCAACCTTATATCAGAAATCCATCTATAGTATC 27513
 QY 2002 ----- 2001
 Db 27514 TACCAAGAAAAAGATCATTTATTTGAAAAAGACCTTGATACAGTATGTTATAGCA 27573
 QY 2002 ----- 2001
 Db 27574 GCACATTCACAAATTTGCAAACTGTGAAACTTAACCAATGTCCATCAGTCAATTAAGTG 27633
 QY 2002 -TAAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2060
 Db 27634 ATTAAGAACTGTGTGTACAGATATATTAATTAATTAATTAATTAATTAATTAATTA 27693
 QY 2061 AATTAATTAATGATGATTCACAGCAATCTAGATGGAATTTGAGACCTTATTTCTAAGTGG 2120
 Db 27694 AATGAATTAACAGATTTGCAAGTACCTGATGATGATTTGAGACCTAATTAATTTCTAAGTGA 27753
 QY 2121 GTAACTCAGGAATGAAAAACCAACATCATATGTTTCTCATTTCAAGTGGGGCTTAAGCT 2180
 Db 27754 GTAACTCAGGAATGAAAAACCAACATCATATGTTTCTCATGATATGAGGATCTTAAGCT 27813
 QY 2181 GTGAGGACAGGAAGCAAT-AGAATGATTAATTAATGAACTGTGGGACTTGAAGGGAAGATG 2239
 Db 27814 ATGAGGACAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 27873
 QY 2240 GAAAGAGGCGAGGATTAAGAACTAGCAATGAGTACAGTGTACATGCTCAGGTGATG 2299
 Db 27874 GAGAGGCGGATGAGGATTAAGAACTAGCAATTAATTAATTAATTAATTAATTAATTA 27933
 QY 2300 GGTGCAACCAAAATCTCAGAAATTAACACTAAGAACTTATTCATGAGAAACA--ACACA 2357

RESULT 8
ADN97989_13
Continuation (14 of 17) of ADN97989 from base 1300001 (Human phosphodiesterase 4D genomic region)

Query Match	16.0%;	Score 1375.6;	DB 12;	Length 110000;
Best Local Similarity	77.1%;	Pred. No. 4.7e-309;		
Matches 1993;	Conservative	0;	Mismatches 409;	Indels 182;
				Gaps 19

OY	1	TGTTATGAAGCCAAATGTCACCTTTTAATACCAAAACCGAAGAAAGATATC - CAAAAAGAA	58
Db	25551	TCATTGAAGCCAGCATCAACCTTGATACCAAAACGAGAAAGGACATTACCAAAAAAGAA	2561
OY	59	ACTATTGACCGAGTACCACTGATGAAATTATACATGACAGAAATCCCAACAAATACTAGCTA	118
Db	25611	ACTACAGATCAATATCTTGATTAACATAGATGTAAATTCCTTAACAAATACCTAGCTG	2567
OY	119	ACCCAATCCACAGCATATCAAGAGATTAATCCACATTTGTCAAGTGGGTTTCATACAG	178
Db	25671	ACTGAATCCAAACAATATCAAAAAGATTAATCCACCATGATCAAGTGGGTTTCATACAG	2573
OY	179	GGGTGCGAG -GATAGGTTAAATACATACACAATCAATTAATGTGATACATCAATAACGAA	237
Db	25731	GGGTGCGAGAGTGGTTTAATGTACACAAGTCAATTAATGTGATACACATATAACGAA	2579
OY	238	TTAAAACAAAAAATCATCATGATCATCTCATAGTGTGTAAAAAGCATTTGACAAAACT	297
Db	25791	TTAAAACAAAAAATTCATGATCATCTCAATAGTGTGAGAAAAAGCATTAACAAAAATCC	2585
OY	298	AACATTTCTTTATGATTAACAACTTCAGCAAAATCGACATAGAAAGGACATACCTTAATG	357
Db	25851	AGCATCCCTTTATGATTTAAAGCTCTCAGCAAAATCAGCATACAGGAGCATACATTAATG	2591
OY	358	TAAATAAAGCCATATATGACGGAACCCACAGCAAAACATTAATACTGAATGGGAAAAGTTGA	417
Db	25911	TAAATAAACATATCTATGACAAACCCACAGCCAAACGTAATAGTAATGGGAAAAGTTGA	2597
OY	418	AAACATTTGTCCTTGAGAACTGGAACAAGACAGAGATG -CTACTTTCAACCACTTCTATTCA	476
Db	25971	AAGAAATTCCTCTGAGAACTGGAACAAGATGATGCCACTCTCACCACTCTTCTTCA	2603
OY	477	ACATAGTAGTGAAGTTTATGACGAGCATTCAGACAGAGAAAGAAATCAAGGGACCC	536
Db	26031	ACATAGTAAGTGAAGTCTTAGCAAGAGCAATTCAGACAGAGGAGAAATTAAGGGACATCC	2609
OY	537	AAATCAATTAAGAGGAAGTCAAACTGTCCCTGTTCACTGATGATATGATTTATACCTAG	596
Db	26091	AAATCGGTAAAGAGGAAGTCAAACTGTCACTGTTTGTGTATGATATATTTTACCTTG	2615
OY	597	AAAAACCTTAAGACTCATCCAGAAAGCTCCTAGAACATGATACATAAATTCAGTAAAGTT	656
Db	26151	AAAACTTAGAACTCCTCAGAGAGCTCCTAGAACATGATTAATGATTAACAAGAGTTT	2621
OY	657	CAGATTCAGAACTTAAATGTACAAATCATGTAGCACTGCTATACACCAACAGTACCAAG	716
Db	26211	CTGATATCAAGATTTAAATGTACAAAAATCATGTACTCTTCTTATATACCAACAGTACCAAG	2627
OY	717	CTGAGATCAAAATCAAGACTCAAAACCTTTTACCAATAGTGTAA ------AAA	765

Db 26271 GGGAGATTCATCAAGAACTCAACCCATTTTACATAGCTGTAAATAATAATAA 26330
 Qy 766 ATACTTAAGAAATATTTCTTACCAAGAGGTGAAGACCTCTACAGAGAAATCAAAAC 825
 Db 26331 ATACTTAAGAAATATTTCTTACCAAGAGGTGAAGACCTCTACAGAGAAATCAAAAC 26390
 Qy 826 ACAGCTGA-----CATCATAGATGACCAAAAGTGTGAACATATCCATGCTCATGAT 881
 Db 26391 ACTGCTGAAGAGAAATCATAGACATATCAAACTAATGGAGACATCCCTGATCATGAT 26450
 Qy 882 GGGTGAATCAATATTTGTGAATAATGACATATGCGCAAAAGCAATCTACAGTTCAATG 941
 Db 26451 GGGTGAATCAATATTTGTGAATAATGCGCATCTGCGCAAAAGCAATCTACAGTTCA 26510
 Qy 942 AATTCACCAAAATATCATCATCTTCTTACAGAACTAGAAAAAAACAATTTCAAAAT 1001
 Db 26511 AATCCCATCAAAATATCAACCATCTTCTTACAGAAATAG-AAAAACAATTTCAAAAC 26569
 Qy 1002 TCATATGGAACAACAACCAAAAAAATAAATACCCGATGCGCAAGCAAGCATTTAG 1061
 Db 26570 TCAGATGGAAC-----CAAAAAAGAGCGCTGATAGCCAAAGCAAGCATTTAG 26615
 Qy 1062 CAAAAAGAACAATCTGAGGAGCATCATTTACCATCTTCAAACTATCTACAGGCTAT 1121
 Db 26616 CAAAAAGTACAAATCTGAGGAGCATCATCTGATTTCAATATATCTATAGGCGCAT 26675
 Qy 1122 AATCACCAAAATCATCTGAGCATGACATTAATAGGACATAGCCATAGAAAGAG 1181
 Db 26676 AGTCCACCAAAATGCGATGCTGTATGTATTAATAATAGCATATAGCAAAATGAAACGAAT 26735
 Qy 1182 AGAGATTCAGAAATATAGCAAAATATATAGCACTGATTTTGAACAAGCAAAACA 1241
 Db 26736 AGAGAAACCGGAGATTAACCCCAATATCTTACAGCAATGATCTTCAAAAGCAAAACA 26795
 Qy 1242 AAAACATTAAGTGGGAAAAA-GACATTTCTAGTTAAACAATGTGTCTGATTTATTTG 1300
 Db 26796 AAAACATTAAGTGGGAAAAAGATTAACCTTTTCAAAATGTGTCTGATTAATTTGG 26855
 Qy 1301 CCACATGTGGAAGATGAAGACCTGATCCCTTGTCTCTACTTAATACAAAAATTTGAT 1360
 Db 26856 CCACACATATGGAAGATGAAGACCTGATCC--TATCTCTACCCGTATCAAAAAATCA 26913
 Qy 1361 AGATGATCAAAAGACTTAATCTGAGACCTTAACCAATAAAATTTCTAGAGATTAACAT 1420
 Db 26914 AGATGATTAAGGGCTTAACCTTAAGACCTGAACTAT-GAAATTTTGAAGATTAATTT 26972
 Qy 1421 AGAAAAATGCTTCTAGACATTTCACTTAGCAAAAGACTTTCAGGCAAGAACCCAAAA 1480
 Db 26973 GAAAAAACCTTCTAGACATTTGCTTAGGCAAGATTTTCAAGCAAGAACCCAAAAAG 27032
 Qy 1481 AATGCACAAAACCAAAATTAATATAGATGAGACTTAATTAAC-TAAAAAGCTTTTGGC 1539
 Db 27033 AATGCAATTAACCAAAAGATTAATATAGCTGGACCTTAATTAACGAGCTTTTGGC 27092
 Qy 1540 AGCAAAAACAAATCATATGAGAGCAAAAGCAAAACCAACCAAGAGTGAAG-AAAACTT 1598
 Db 27093 GCAAAAAGAGACATGAGAGATTAACAGCAAAACCAACCAAGAGTGAAGAAAAATCTT 27152
 Qy 1599 CAATCTATATCTCTGCAAAAGG-GTAGTATCCAGAAATCTTCAAGAGAACCCAAACA 1658
 Db 27153 CAATCTATATCTCTGCAAAAGG-GTAGTATCCAGAAATCTTCAAGAGAACCCAAACA 27211
 Qy 1659 GCAAGAGAGAGCAAAATCTCCATGAAAGAGTGGCTAAGAGACATGATAGCAATTT 1718
 Db 27212 GTAAGAAAAAAACCAAACTATCCCATCAAAAGTAAAGGCTAAGAGAGCAATTTCA 27271
 Qy 1719 CAAAAGAGATATACAAATGAGCAAAAC--AGAAAAAATGCTTAACTACTAATGA 1776
 Db 27272 CAAAAGAGATATACAAATGAGCAAAACATATGAAAAAATGCTCAACATCTAATGA 27331
 Qy 1777 TTAGGAAATGTAAATCAACA--CTGTATGCAATACCACTTACTCTGCAAGAAATGT 1834

Db 27332 TCAGAGAAATGCMAATCAAAACACAAATGTGATACACCGTACTTCTGCAAGATG 27391
 Qy 1835 CATTAATTTAAATCT-AAAAATATAGATGTTGGTGGTCTGTGCTGATTAAGAAACA 1892
 Db 27392 CATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 27451
 Qy 1893 CTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1952
 Db 27452 CTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 27511
 Qy 1953 TTTCTTAAAGAACTAAAGTATGATGACCAATTTGATCCAGCAATCCAT----- 2001
 Db 27512 TTTCTTAAAGAACTAAAGTATGATGACCAATTTGATCCAGCAATCCAT----- 27571
 Qy 2002 ----- 2001
 Db 27572 TACCCAGAGAAAGAAAGTATTTTGAAGAAACAATTGATACGATATGTTATGCA 27631
 Qy 2002 ----- 2001
 Db 27632 GCACAAATTCACAAATTCGCAAACTGTGGAACCTAACCCAAATGTCATCAATTA 27691
 Qy 2002 -TAAATATGTATTAATATATATATATATATATATATATATATATATATATAT 2060
 Db 27692 ATTAAGAACTGTGTGTACAGATATATATATATATATATATATATATATATAT 27751
 Qy 2061 AATTAAT 2120
 Db 27752 AATGAATTAACAGCATTTGACGATGACCTGATGATGATGATGATGATGATG 27811
 Qy 2121 GTAACTCAGGAATGGAAGAAACCAAACTATATGTTCTCACTTAACAAGTGGGCTAAG 2180
 Db 27812 GTAACTCAGGAATGGAAGAAACCAAACTATATGTTCTCACTTAATGTTGATCT 27871
 Qy 2181 GTGAGACACGAAGGAT-AGAAATGATATATATATATATATATATATATATAT 2239
 Db 27872 ATGAGACACGAAGGATATATATATATATATATATATATATATATATATAT 27931
 Qy 2240 GAAGAGAGCGAGGATTAATATATATATATATATATATATATATATATATATAT 2299
 Db 27932 GAGGAGGAGGATGAGGATTAATATATATATATATATATATATATATATAT 27991
 Qy 2300 GGTGACACCAAAATCTCAGAAATTTACATTAAGAACTTATCCATGAGAGCA--AC 2357
 Db 27992 GGTGACACCAAAATCTCAGAAATTTACATTAAGAACTTATCCATGAGAGCA 28051
 Qy 2358 CTTGTTCCCAAAATCCCAATGAAATTAATATATATATATATATATATATAT 2417
 Db 28052 CTTGTTCCCAAAATCCCAATGAAATTTTATATATATATATATATATATAT 28111
 Qy 2418 AATT 2421
 Db 28112 GTTT 28115

RESULT 9
 AAD33751
 ID AAD33751 standard; DNA; 1400 BP.
 XX AAD33751;
 AC
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human MrgX3 (mas-related gene) DNA.
 XX
 KM Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
 XX receptor; sensory perception; pain; analgesic; MrgX3; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 332..1300
 FT /*tag= a

DE Human G protein-coupled receptor MrgX3 nucleotide seq ID NO:673.
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related disease; cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer; gene; ds.
 XX Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2001; 2000US-0257144P.
 PR
 XX 19-DEC-2000; 2000US-0257144P.
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX Burner GC, Roush CL, Brown JP;
 FI
 XX WPI, 2003-046718/04.
 DR P-PSDB; ABB81750.
 XX
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 PT
 XX Disclosure; Fig 1; 523bp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABB82019 to ABB83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcer, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABB82523 to ABB82869 encode
 CC GPCR proteins given in ABB81675 to ABB82018, which are used in the
 CC exemplification of the present invention
 CC
 XX
 XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;
 SQ
 Query Match 15.9%; Score 1374.6; DB 8; Length 1400;
 Best Local Similarity 99.3%; Pred. No. 9.1e-310;
 Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 7196 TCAGGCCAGATGATGATATCATCGGGTCCACAGCACTGCTAGATGAGTGGGGCTTT 7255
 DB 1 TCAGGCCAGATGATGATATCATCGGGTCCACAGCACTGCTAGATGAGTGGGGCTTT 60

QY 7256 TTGATCCTAATGTTATCCCATGTGACAGACAGAACTTGNGGSCAGTAGAGAGTCCAG 7315
 DB 61 TTGATCCTAATGTTATCCCATGTGACAGACAGAACTTGNGGSCAGTAGAGAGTCCAG 120
 QY 7316 GCTTCAGAGTCAACAGAACTGGATTTCAACTGAGATTAGAGACCCCACTTTTGATA 7375
 DB 121 GCTTCAGAGTCAACAGAACTGGATTTCAACTGAGATTAGAGACCCCACTTTTGATA 180
 QY 7376 GGTGACTTATTTCTTCGAGGCTCTGATCTCTCTTTTAAATGAGACAGTAAATCCCA 7435
 DB 181 GGTGACTTATTTCTTCGAGGCTCTGATCTCTCTTTTAAATGAGAGAGTAAATCCCA 239
 QY 7436 CATGGAGGGGTGGGGAGATCAGAGATCAAAACAGCTGATCAGATCTGGTTTCTG 7495
 DB 240 CATGGAGGGGTGGGGAGATCAGAGATCAAAACAGCTGATCAGATCTGGTTTCTG 299
 QY 7496 TTTCAGGGTCAATCAGACTGGGGTTTCTAGAGATGAGATTCAACATCCAGTCTTGAGTA 7555
 DB 300 TTTCAGGGTCAATCAGACTGGGGTTTCTAGAGATGAGATTCAACATCCAGTCTTGAGTA 359
 QY 7556 CAGACTGACCAATCAACAGAGAGTGAAGAGACTCTTGTCTTCAACAGACAGACCTGAGCT 7615
 DB 360 CAGACTGACCAATCAACAGAGAGTGAAGAGACTCTTGTCTTCAACAGACAGACCTGAGCT 419
 QY 7616 TCAGGGGGGTGACGTGATCGTTTCCCTGTGGCGCTGACAGGAAACGGGTTGCTCT 7675
 DB 420 TCAGGGGGGTGACGTGATCGTTTCCCTGTGGCGCTGACAGGAAACGGGTTGCTCT 479
 QY 7676 GGTCTCTGGGCTGACGTGATCGGAGAGACGTGTCTCATCTTCAATCTTCAACCTGATCG 7735
 DB 480 GGTCTCTGGGCTGACGTGATCGGAGAGACGTGTCTCATCTTCAATCTTCAACCTGATCG 539
 QY 7736 CGGCGCACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 7795
 DB 540 CGGCGCACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 599
 QY 7796 TCCGCAATCCCATCTCAAAATCTCAGTCTGTGATGACCTTCCCTTCTTCTTCTTCTTCT 7855
 DB 600 TCCGCAATCCCATCTCAAAATCTCAGTCTGTGATGACCTTCCCTTCTTCTTCTTCTTCT 659
 QY 7856 TAAAGATGCTGAGCGGCATCAGACAGGAGCGTCTGTCATCTCTGAGGCCATCTGCT 7915
 DB 660 TAAAGATGCTGAGCGGCATCAGACAGGAGCGTCTGTCATCTCTGAGGCCATCTGCT 719
 QY 7916 ACCACTGCGCGCCGCCCAATACCTGTCAATGCTATGCTGCTCTGCTCTGCGCCCTGT 7975
 DB 720 ACCACTGCGCGCCGCCCAATACCTGTCAATGCTATGCTGCTCTGCTCTGCGCCCTGT 779
 QY 7976 CCGTGCAGGAGATCTCTGAGTGTGATGCTGACTTCTGTTAGTGTGATG 8035
 DB 780 CCGTGCAGGAGATCTCTGAGTGTGATGCTGACTTCTGTTAGTGTGATG 839
 QY 8036 CTGTTTGTGTGAAGAGTCAATGATTCATTAATCGCGTGTGATTTTATATGTGTG 8095
 DB 840 CTGTTTGTGTGAAGAGTCAATGATTCATTAATCGCGTGTGATTTTATATGTGTG 899
 QY 8096 TTCTGTGTGTGAGTCACTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8155
 DB 900 TTCTGTGTGTGAGTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 959
 QY 8156 CGCTGACAGGCTGTGAGTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8215
 DB 960 CGCTGACAGGCTGTGAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1019
 QY 8216 TGCCCTTTGAGTCACTGT 8275
 DB 1020 TGCCCTTTGAGTCACTGT 1079
 QY 8276 TTGTGATGTGATCTAGT 8335
 DB 1080 TTGTGATGTGATCTAGT 1139

QY 8336 TCATTATCTTCTGTTGGGCTCTTTAGGACGCTCAAAATAGGACGAACTCTGAAGCTGG 8335
 DB 1140 TCATTATCTTCTGTTGGGCTCTTTAGGACGCTCAAAATAGGACGAACTCTGAAGCTGG 1139
 QY 8396 TTCTCCAGAGGGCTCTGACAGACAGCCTGAGTGAAGTGAAGTGGGCTTCTC 8455
 DB 1200 TTCTCCAGAGGGCTCTGACAGACAGCCTGAGTGAAGTGAAGTGGGCTTCTC 1239
 QY 8456 AGGAAACCTTGAGCTGTGCGGAAAGCAGATTGAGCAGTGAAGAACTCTGCTGT 8515
 DB 1260 AGGAAACCTTGAGCTGTGCGGAAAGCAGATTGAGCAGTGAAGAACTCTGCTGT 1319
 QY 8516 CAGACAGACTTGAAGCAATGCTGCTGCGACCTTGAACATTATATGATTTTCT 8575
 DB 1320 CAGACAGACTTGAAGCAATGCTGCTGCGACCTTGAACATTATATGATTTTCT 1379
 QY 8576 TAGCCTTCTGCTCAGAAATG 8596
 DB 1380 TAGCCTTCTGCTCAGAAATG 1400

RESULT 11
 ADH08534
 ID ADH08534 standard; DNA; 1400 BP.
 XX
 AC ADH08534;
 XX
 DT 25-MAR-2004 (first entry)
 XX
 DE DNA sequence MrgX3.
 XX
 KW mas-related gene D; MrgD; Analgesic; Vulnerary; Ophthalmological;
 KM sensory perception; glaucoma; Mrg; ds.
 OS Mus musculus.
 XX
 PN W02004003133-A1.
 XX
 PD 08-JAN-2004.
 XX
 PF 13-MAY-2003; 2003MO-US015004.
 XX
 PR 26-JUN-2002; 2002US-00183116.
 PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 PI Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
 XX
 DR WPI; 2004-083025/08.
 XX
 DR P-PSDB; ADH08535.
 PT New mas-related gene D polypeptides, useful as therapeutics or in
 PT identifying agonists or antagonists that alter pain perception in a
 PT mammal for treating impaired sensory perception, e.g. chronic intractable
 PT pain or neuropathic pain.
 XX
 PS Disclosure; SEQ ID NO 30; 220bp; English.
 XX
 CC The present invention relates to an isolated mas-related gene D (MrgD)
 CC polypeptide. The MrgD polypeptides are useful as therapeutics or for
 CC identifying compounds, i.e. agonists or antagonists, that alter pain
 CC perception in a mammal. The compounds are useful for treating impaired
 CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
 CC promoting wound healing, restoring normal sensitivity following injury,
 CC or treating ocular conditions, particularly those associated with
 CC pressure such as glaucoma. The Mrg genes or proteins may be used as
 CC molecular probes for the detection of cells or tissues related to or
 CC involved with sensory perception. The present sequence represents a MrgA
 CC (Mrg subfamily) encoding sequence.
 XX
 SQ Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;
 Query Match 15.9%; Score 1374.6; DB 12; Length 1400;

Best Local Similarity 99.3%; Pred. No. 9,1e-310;
 Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 7196 TCAGGCCAGAGATAGTAATCATCGGGTCCACAGACATGGCTAGATGAGTGGGGTGT 7255
 DB 1 TCAGGCCAGAGATAGTAATCATCGGGTCCACAGACATGGCTAGATGAGTGGGGTGT 60
 QY 7256 TTGATCTTAATGTTATCCCATGTCAGACAGAACTGTGTGGCAGTGAAGAGTCAAG 7315
 DB 61 TTGATCTTAATGTTATCCCATGTTAGACAGAACTGTGTGGCAGTGAAGAGTCAAG 120
 QY 7316 GCTTCAGAGTCAAGAAAGCTGAAATTTCAACTGGAATTTAGAACCCCACTTTGATA 7375
 DB 121 GCTTCAGAGTCAAGAAAGCTGAAATTTCAACTGGAATTTAGAACCCCACTTTGATA 180
 QY 7376 GGTGACTTATTTCTGCGAGCTCTGATCTCTCTCTTTAATGAGACAGTAAATCCCA 7435
 DB 181 GGTGACTTATTTCTGAGAGTCTGATCTCTCTCTTTAATGAGACAGTAAATCCCA 239
 QY 7436 CATGGCAGGGTGTGGGAGATCAAGATCAAAACGTGGTATCATCTGTTCTG 7495
 DB 240 CATGGCAGGGTGTGGGAGATCAAGATCAAAACGTGGTATCATCTGTTCTG 299
 QY 7496 TTTCAGGGTATCAGACTGGGGTTTCTGAGCATGATTTCAACCATCCAGTCTGGATA 7555
 DB 300 TTTCAGGGTATCAGACTGGGGTTTCTGAGCATGATTTCAACCATCCAGTCTGGATA 359
 QY 7556 CAGAACTGACACCAATCAACGAGTGAAGAGACTCTTGTGTAAGAGACACCTGAGCT 7615
 DB 360 CAGAACTGACACCAATCAACGAGTGAAGAGACTCTTGTGTAAGAGACACCTGAGCT 419
 QY 7616 TCACGGGGCTGACGTGATCGTTTCCCTGTGCGCTGACAGAAACGGGTGTGCTCT 7675
 DB 420 TCACGGGGCTGACGTGATCGTTTCCCTGTGCGCTGACAGAAACGGGTGTGCTCT 479
 QY 7676 GGTCTCTGGGGTGGCGGATGGGAGAAAGCTGTCTCATCTAATCTCTCAACCTGTGCG 7735
 DB 480 GGTCTCTGGGGTGGCGGATGGGAGAAAGCTGTCTCATCTAATCTCTCAACCTGTGCG 539
 QY 7736 CGGCCGACTTCTCTCTTCTTGGCGGACATTTATATGTTGCGCTTACGCTCATCAATA 7795
 DB 540 CGGCCGACTTCTCTCTTCTTGGCGGACATTTATATGTTGCGCTTACGCTCATCAATA 599
 QY 7796 TCGGCATCCCATCTCCAAAATCTCACTGTGTGATGACCTTTCCCTATTATAGGCC 7855
 DB 600 TCGGCATCCCATCTCCAAAATCTCACTGTGTGATGACCTTTCCCTATTATAGGCC 659
 QY 7856 TAAAGCATGTAAGGGCCATGACGACCGAGCGCTGCTGCATCTGTGGGCCATCTGGT 7915
 DB 660 TAAAGCATGTAAGGGCCATGACGACCGAGCGCTGCTGCATCTGTGGGCCATCTGGT 719
 QY 7916 ACCACTGCGCGGCCAGATACCTGTCACTGATGTGATGTCTGTGCGGCCCTGT 7975
 DB 720 ACCACTGCGCGGCCAGATACCTGTCACTGATGTGATGTCTGTGCGGCCCTGT 779
 QY 7976 CCTGTGCGGAGTATCTTGAGTGTGATGTTCTGTGACCTTCTGTTAATGATGTGAT 8035
 DB 780 CCTGTGCGGAGTATCTTGAGTGTGATGTTCTGTGACCTTCTGTTAATGATGTGAT 839
 QY 8036 CTGTTTGTGTGAAGACGTCAATTTCAATTAACCGGTGTGCTGTTTATATGTGCG 8095
 DB 840 CTGTTTGTGTGAAGACGTCAATTTCAATTAACCGGTGTGCTGTTTATATGTGCG 899
 QY 8096 TTCTCTGTGGGTCAAGCTGTGCTGTGATGAGATTTCTGTGATCCCGAAGATGC 8155
 DB 900 TTCTCTGTGGGTCAAGCTGTGCTGTGATGAGATTTCTGTGATCCCGAAGATGC 959
 QY 8156 CGCTGACAGGCTGATGATGATCAATCTCTGACAGTGTGATCTCTCTCTGTGCGC 8215
 DB 960 CGCTGACAGGCTGATGATGATCAATCTCTGACAGTGTGATCTCTCTCTGTGCGC 1019
 QY 8216 TGCCCTTTGATTCATGATGGGCGCTGTTTCCAGATTCACCTGATTTGAAGAAATCTTAT 8275

```

Db      1020 TGCCCTTGGCATTCAGTGGGCCCTGTTTCCAGATTCACCTGATGGAAAGTCTTAT 1079
Qy      8276 TTTGTGATGTCATCTAGTTTCCATTTCTGTCCGCTCTTAAACAGAGTCCCAACCCA 8335
Db      1080 TTTGTGATGTCATCTAGTTTCCATTTCTGTCCGCTCTTAAACAGAGTCCCAACCCA 1139
Qy      8336 TCATTACTCTTCTGCGGCTCTTTAGGAGGTCGCAAAATGGCGAAACCTGAGCTGG 8395
Db      1140 TCATTACTCTTCTGCGGCTCTTTAGGAGGTCGCAAAATGGCGAAACCTGAGCTGG 1199
Qy      8396 TTTCTCAGAGGGCTCTGACAGACACGCTGAGGTGATGATGAGGTGAGGCTTCTC 8455
Db      1200 TTTCTCAGAGGGCTCTGACAGACACGCTGAGGTGATGATGAGGTGAGGCTTCTC 1259
Qy      8456 AGAAACCTTGAGCTGTGCGGAGAGCAATTTGAGAGTGAAGAACTTCTGCTCT 8515
Db      1260 AGAAACCTTGAGCTGTGCGGAGAGCAATTTGAGAGTGAAGAACTTCTGCTCT 1319
Qy      8516 CAGACAGACCTTGAAGAGAAAGCTGCCGACACCTTGACAAATTAATGCAATTTTCT 8575
Db      1320 CAGACAGACCTTGAAGAGAAAGCTGCCGACACCTTGACAAATTAATGCAATTTTCT 1379
Qy      8576 TAGCCTTCTGCTCAGAAATG 8596
Db      1380 TAGCCTTCTGCTCAGAAATG 1400

RESULT 12
AD050281_13
Continuation (14 of 17) of AD050281 from base 1300001 (Human phosphodiesterase 4D (PDE4D
WP Sequence split into 17 fragments LOCUS AD050281 Accession AD050281
WP      Fragment Name      Begin      End
WP      AD050281_00          1      110000
WP      AD050281_01          100001   210000
WP      AD050281_02          200001   310000
WP      AD050281_03          300001   410000
WP      AD050281_04          400001   510000
WP      AD050281_05          500001   610000
WP      AD050281_06          600001   710000
WP      AD050281_07          700001   810000
WP      AD050281_08          800001   910000
WP      AD050281_09          900001  1010000
WP      AD050281_10         1000001  1110000
WP      AD050281_11         1100001  1210000
WP      AD050281_12         1200001  1310000
WP      AD050281_13         1300001  1410000
WP      AD050281_14         1400001  1510000
WP      AD050281_15         1500001  1610000
WP      AD050281_16         1600001  1691134

Query Match      15.9%; Score 1374; DB 12; Length 110000;
Best Local Similarity 77.1%; Pred. No. 1.1e-308;
Matches 1992; Conservative 0; Mismatches 410; Indels 182; Gaps 19;

```

```

Qy      298 AACATTTCTTATGATTAACCTTCAGCAAAATGCAATGAAAGACATACCTTAATG 357
Db      25850 AGATCCCTTTATGATTAAGCTCTCAGCAAAATGACATACAGGACATACATTAATG 25909
Qy      358 TAATAAAGCCATATATGACGACCCACAGCAAAATTAATGATGAGGAAAAAGTTGA 417
Db      25910 TAATAAAGCCATATATGACCAACCAAGCCAAAGTAATGATGAGGAAAAAGTTGA 25969
Qy      418 AACATTTCTTATGATTAACCTTCAGCAAAATGCAATGAAAGACATACCTTAATG 476
Db      25970 AAGATTTCTTATGATTAAGCTCTCAGCAAAATGACATACAGGACATACATTAATG 26029
Qy      477 ACATAGTAGGAAAGTTTAAAGCCATGACAAAGCAAAAGAAAGAAATCAAGGACACC 536
Db      26030 ACATAGTAGGAAAGTTTAAAGCTCTCAGCAAAAGCAATGACAAAGGAAAGAAATTAAGGACATC 26089
Qy      537 AATCAATTAAGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATGATGATGATGATG 596
Db      26090 AATCGTAAAGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATGATGATGATGATGATG 26149
Qy      597 AAAACCTTAAAGACTATCCAGAAAGCTCTGAAATCTGATATATATTAATTCAGTAAGTT 656
Db      26150 AAAACCTTAAAGACTATCCAGCAAGCTCTGAAATCTGATATATTAATTCAGTAAGTT 26209
Qy      657 CAGATTAACAACTAATGATGACAAATGATGACATGCTGATATACCAACAGTACCAAG 716
Db      26210 CTGATTAACAACTAATGATGATGACAAATGATGACATGCTGATATATACCAACAGTACCAAG 26269
Qy      717 CTGATTAACAACTAATGATGACAAATGATGACATGCTGATATATACCAACAGTACCAAG 765
Db      26270 GGGAGATTAACAACTAATGATGACAAATGATGACATGCTGATATATACCAACAGTACCAAG 26329
Qy      766 ATATCTTAAGATATTTCTTACCAAGAGTGAAGACCTCTCAAGGAAATTAATTAATTAAC 825
Db      26330 ATATCTTAAGATATTTCTTACCAAGAGTGAAGACCTCTCAAGGAAATTAATTAATTAAC 26389
Qy      826 ACAGCTGA-----CATCATGATGACACAAACAAAGTGAAGACATCCATGCTCATGAT 881
Db      26390 ACTGCTGAAGAGATCATGACATGACATGACAAATGGAACACATCCGTCATGAT 26449
Qy      882 GGGTAGATCATATATTTGTAAGAAATGACATATTTGCCAAAGCAATCTCAAGTCAATGC 941
Db      26450 GGGTAGATCATATATTTGTAAGAAATGACATATTTGCCCAATCTCAAGTCAATGC 26509
Qy      942 AATTTCCACCAAAATATCATCATATTTCTTACAGAACTGAAAGAAACAAATTTCTTAAT 1001
Db      26510 AATTTCCACCAAAATATCATCATATTTCTTACAGAACTGAAAGAAACAAATTTCTTAAT 26568
Qy      1002 TCATATGGAACCAACCAAAAAAAAAAAAAAAAAAACCAGATAGCAAGCAAGACTTAG 1061
Db      26569 TCAGATGGAAC-----CAAAAGAGCTGCTGACCAAGCAAGCAAG 26614
Qy      1062 CAAAAAGAACAAATCTGAGAGCATCAATTAATCCATCTTCAAACTATCAAGGCTAT 1121
Db      26615 CAAAAAGTAACAAATCTGAGAGCATCAATTAATCCATCTTCAAACTATCAATTAATG 26674
Qy      1122 AATTCACCAAAATCATGAGCATGATTAATCTAGGACATGACATGACATGACATGACATG 1181
Db      26675 AGTCACCAAAATGAGCATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 26724
Qy      1182 AGAAATTCAGAAATTAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1241
Db      26735 AGAAATTCAGAAATTAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26794
Qy      1242 AAACATTAAGTGGGAAAAA-GACATTTCTAGTTTAACAAATGCTGAGATTAATTTGCAAG 1300
Db      26795 AAACATTAAGTGGGAAAAAAGATTAATTTTCAACAAATGCTGAGATTAATTTGCTAG 26854
Qy      1301 CCACATGGAAGAAATGAATCTGATCCCTTGTCTCACTTAATTAACAAATTAATTAATTAAT 1360
Db      26855 CCACATGGAAGAAATGAATCTGATCCCTTGTCTCACTTAATTAACAAATTAATTAATTAAT 26912

```

OY	1361	AGATGATCAAAAGACTTAATCTGAGACCTTAAACCTATAAAATCTAGAAAGTAACTC	1420
Db	26913	AGATGATTTAAGGCTTAACTTAAGCGTGAACCTAT -GAATTTTGAAGATTACTTT	26971
OY	1421	AGAAAAATGCTTGAACATTCACTTAGGCAAGACTTCAGGCAAGAACTCCAAAGTA	1480
Db	26972	GGAAAAACCCCTTGAACATTTGGCTTGGCAAGGATTTCTGACCAAGAACTCCAAAGTA	27031
OY	1481	AATCCACAAAACAAAATTAATATAGTATAGACTTAATTAAC -TAAAAAGCTTTTGGCC	1539
Db	27032	AATCCATTAATAAACAAAGATTAATATAGTATAGGACTTCACTTAATTAACGAGCTTTTGCAG	27091
OY	1540	AGCAAAAACATCAATTGACAGCAAAACAGCAACCAACGAGTGAAG -AAAATCTTCA	1598
Db	27092	GGCAAAAGAAACAGTCAAGCAAGATTAACAGCAACCAACAGATGGAGAAATCTTCA	27151
OY	1599	CAAACTAAGACTTGACTAAGAACTAATATCCGAACTCCAAAGAACTCAAACAATCA	1658
Db	27152	CAATCTAATCTCTGACAAAGG -GTACTATCCAGAAATCTACAAAGAACCCAAACAATCA	27210
OY	1659	GCAAGAAAGAACCAACATCCCATGAAAGTGGCTTAAGACATGATATGCAATTCT	1718
Db	27211	GTAAGAAAAAAACAAACAATCCCATCAAAAAGTAGGCTAAGGCAATGAGCAATTCA	27270
OY	1719	CAAAAAGAAATATCAAAATGGCCAAACAAC -AGSAAAAATGCTTAACATCACTAATGA	1776
Db	27271	CAAAAAGAAATATCAAAATGGCCAAACAACATTAAGAAAAATCTCAACATCACTAATGA	27330
OY	1777	TTAGGAAATGTAATATCAACA -CTGATATGCGATACACCTTATCTCTCAAGAAATGT	1834
Db	27331	TCAAGAAATGCAAAATCAAAACCAAAATATGATATACCAACGATCTTCTGCAAGAAATGC	27390
OY	1835	CATAATTTAAAAATCT -AAAAATATAGATGTTGTGGTCTGTGCTGATTAAGAAACA	1892
Db	27391	CATAATTTAAAAATTTTAAAAAACAGATAGATGTTGGCATGAAAGGGGTGATCAGAAACA	27450
OY	1893	CTTTTACACGCTGCTGAGAAATGTAACCTTGCCGAACCACTATAGAAAAACAGTGTGAAA	1952
Db	27451	CTTCTACACGCTGCTGAGAAATGTAACCTTGCCGAACCACTATAGAAAAACAGTGTGAAA	27510
OY	1953	TTTCTTAAGAACTTAAGATAGATCGACCAATTTGATCAGACCAATCCCAT -	2001
Db	27511	TTCTTTAAGAACTTAAGATAGATCGACCAATTTGATCAGACCAATCCCAT -	27570
OY	2002	-----	2001
Db	27571	TACCCAGAGAAAAAGATCATTTATTTGAAAAAGACCTTGTAACGATATTTATAGCA	27630
OY	2002	-----	2001
Db	27631	GCACAATTCACAAATTCGAAAACTGTGAACTTAACCCAAATGTCCATCATTAAGTGA	27690
OY	2002	-TAAATATGTAATTAATATATATTTATATACATGGAATATACAACTCAGCATTAAGAAAG	2060
Db	27691	ATTAAGAACTGTGTGTACAGATATATATACATGGAATATATACAACTCAGCATTAAGAAAG	27750
OY	2061	AATAAAATGATGATTCACAGCAATCTAGATGGAATTTGAGAACCTTATTTCTAAGTGG	2120
Db	27751	AATGAATTAACAGACTTTGACAGTACCTGATGAGATTTGAGACATATTTATTTCTAAGTGA	27810
OY	2121	GTAATCTCAGAAATGAAAAACCAACATCATATGTTCTCACTTACAAAGTGGGGCTAAGCT	2180
Db	27811	GTAATCTCAGAAATGAAAAACCAACATCATATGTTCTCACTTACAAAGTGGGGCTAAGCT	27870
OY	2181	GTAAGGACAGAAAGGACAT -AGAAATGATATATAGAACTGTGGGAGCTTGAAGGGGAAGATG	2239
Db	27871	ATGAGGACACAAAGATATTAAGAAATGATACATGGAATTTTGGGAGCTTGAAGGGGAAGATG	27930
OY	2240	GAAAGAGGCGAGAGGATTAAGACTACACATATGAGTCAAGTGTACACTGCTCAGGTATG	2299
Db	27931	GGAAGGGGGGTGAGGATTAAGAAATTAACAATATGTGTCAAGTATATCTGTTGGGTATG	27990
OY	2300	GGTGACCAAAATCTGAAATTAACCACTTAAGAACTTATCATGGAAGCA -ACACCA	2357

Dd		27991	GATGACCAAAATCTCAGCAATCACCACTTAAGAATTCTTACTATATACTAATAATATACCA	2805
Oy		2358	CTGTTCCTCCCAAAAATGCCAATGAATTAATAATATATAATGATTTAATTTCACG	2417
Dd		28051	CCTGTACCCTTAATACTATTATGAAAAAAATTTTATATAAAGTAAATGATTTAAGTCAGAA	28110
Oy		2418	AATT 2421	
Dd		28111	GTTT 28114	
RESULT 13				
ID	ABN96931/C			
XX	ABN96931 standard; DNA; 99014 BP.			
XX	ABN96931;			
XX	13-AUG-2002 (first entry)			
XX	Gene #3429 used to diagnose liver cancer.			
DE	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;			
KW	metastatic liver tumour; cytostatic; expression profile; disease state;			
KM	disease progression; drug toxicity; drug efficacy; drug metabolism.			
XX	Homo sapiens.			
OS	MO200229103-A2.			
PN	11-APR-2002.			
PD	02-OCT-2001; 2001WC-US030589.			
PF	02-OCT-2000; 2000US-0237054P.			
PR	(GENE-) GENE LOGIC INC.			
PA	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;			
PI	WPI; 2002-426119/45.			
DR				
XX				
PT	Diagnosing and detecting the progression of liver cancer, hepatocellular			
PT	carcinoma or metastatic liver tumor in a patient, involves detecting the			
XX	level of expression of two or more genes in a liver tissue sample.			
PS	Claim 1; SEQ ID NO 3429; 298pp; English.			
XX				
CC	The invention relates to a novel method for diagnosing and detecting the			
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver			
CC	tumour in a patient, and differentiating metastatic liver cancer from			
CC	hepatocellular carcinoma in a patient, involving detecting the level of			
CC	expression of two or more genes represented in ABN93503-ABN97455 in a			
CC	tissue sample. The method of the invention has hepatotropic, and			
CC	cytostatic activity. The method is useful for diagnosing and detecting			
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic			
CC	liver carcinoma in a patient. The method is useful for identifying			
CC	expression profiles which serve as useful diagnostic markers as well as			
CC	markers that can be used to monitor disease states, disease progression,			
CC	drug toxicity, drug efficacy and drug metabolism. Note: The sequence data			
CC	for this patent did not form part of the printed specification, but was			
CC	obtained in electronic format directly from WIPO at			
CC	ftp.wipo.int/pub/published_pct_sequences			
SQ	Sequence 99014 BP; 27179 A; 21575 C; 21336 G; 28924 T; 0 U; 0 Other;			
Query Match	15.8%; Score 1356; DB 6; Length 99014;			
Best Local Similarity	78.6%; Pred. No. 5,6e-305;			
Matches 1895; Conservative	0; Mismatches 345; Indels 170; Gaps 16;			
Oy	131 AGCATATCAAGATATATACCACTTCGCATAGTGATTTCATACAGGGGTGACGA-T 189			
Dd	88976 AGAAAAAATTAAGATATATACCACTTCGTTCACTTGGTTTCATACAGGAAATGACGACT 88917			

QY 190 AGTTAATACATACAGATGATTAATGTGATACATGATTAACAGATTTAAAAA 249
| | | | |
Db 88916 GGTTTAACTAATGCAAGTTAAATGATGATATACACATTAACAGATTTAAAAA 88857
| | | | |
QY 250 ATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTTGAACAATCTAATCTTTA 309
| | | | |
Db 88856 ATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTTGAACAATCTAATCTTTA 88797
| | | | |
QY 310 TGATTTAAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAATAAGCCA 369
| | | | |
Db 88796 TGATTTAAACCTTCAGCAAAATGCGCATACAAAGCATCTCAATGTAATAAGCCA 88737
| | | | |
QY 370 TATATAGCGAGCCCAAGCAAACTATATCTGTAATGGGAAAAAGTTGAAAACATTTGCC 429
| | | | |
Db 88736 TCTATAGCAACCCCAAGCCCAATATATCTGTAATGGTAAAAAGTTGAAAACATTTGCC 88677
| | | | |
QY 430 TGAGAACTGGAACAAGACAGAGATG-CTACTTTCAACAATCTTATCAATAGTAGTG 488
| | | | |
Db 88676 TGAGAACTGGAACAAGACAGAGATGCCACTCTCACTACTCTTCAATAGTAGTG 88617
| | | | |
QY 489 AAGTTTACGAGAGCAATCAGACAGAGAAAGAAATCAAGGCGCAACCAATCAATAG 548
| | | | |
Db 88616 AAGTTTACGAGAGCAATCAGACAGAGAAAGAAATCAAGGCGCATCC-AATTGTPAAG 88558
| | | | |
QY 549 AGGAATGCAACTGTCTCTTTCACTGATGATGATGATGATGATGATGATGATGATG 608
| | | | |
Db 88557 AGGAATGCAACTGTCTCTTTCACTGATGATGATGATGATGATGATGATGATGATG 88498
| | | | |
QY 609 ACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGATACAA 668
| | | | |
Db 88497 ACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGATACAA 88438
| | | | |
QY 669 TAAATGTACAAATCAGTACAGCTGTATACCAACAGTACCAAGCTGAGATCAA 728
| | | | |
Db 88437 TAAATGTACAAATCAGTACAGCTGTATACCAACAGTACCAAGCTGAGATCAA 88378
| | | | |
QY 729 TCAAGAACTCAAACTTTTACAAATAGCTGTAA-----AAATTTCTTAAGATTTCT 782
| | | | |
Db 88377 TCAAGAACTCAAACTTTTACAAATAGCTGTAA-----AAATTTCTTAAGATTTCT 88318
| | | | |
QY 783 TACCCAGAGAGTGAAGACCTCTACAGAGAAAACTACAAACACAGCT-----GACATCA 838
| | | | |
Db 88317 TACCCAGAGAGTGAAGACCTCTACAGAGAAAACTACAAACACAGCTGTGAAGAAAGCA 88258
| | | | |
QY 839 TAGATGACAAACAGTGAAGACATCCATGCTCATGATGATGATGATGATGATGATG 898
| | | | |
Db 88257 TAGATGACAAACAGTGAAGACATCCATGCTCATGATGATGATGATGATGATGATG 88198
| | | | |
QY 899 TGAATAATGACATATGCGCAAAAGCAATCTACAAAGTTCAATGCAATCCCAACAAATAT 958
| | | | |
Db 88197 TGAATAATGACATATGCGCAAAAGCAATCTACAAAGTTCAATGCAATCCCAACAAATAT 88138
| | | | |
QY 959 CATCATCATCTTCACAGAACTAGAAAAAACAATTTAAATTCATATGAAACAACAC 1018
| | | | |
Db 88137 CATCATCATCTTCACAGAACTAGAAAAAACAATTTAAATTCATATGAAACAAC----- 88085
| | | | |
QY 1019 CAAAAAAGGCGCCCAATAGCCAAAGCAAACTAGCAAAAAAACAATCTG 88084
| | | | |
Db 88084 -----CAAAAAAGGCGCCCAATAGCCAAAGCAAACTAGCAAAAAAACAATCTG 88033
| | | | |
QY 1079 GAGGATGACATTAACCATCTTCAAACTATATCAAGGCTATTAATCAACAAACATCAT 1138
| | | | |
Db 88032 GAGGATGACATTAACCATCTTCAAACTATATCAAGGCTATTAATCAACAAACATCAT 87973
| | | | |
QY 1139 GGCATGACATTAACCTAGGACATAGACCAATGGAAGAAAGAGAGATCTCAGAAATTA 1198
| | | | |
Db 87972 GGCATGACATTAACCTAGGACATAGACCAATGGAAGAAAGAGAGATCTCAGAAATTA 87913
| | | | |
QY 1199 AGCCAAATTAATTAAGCAACTGATTTTGAACAAAGCAACAAACATTAAGTGGGAA 1258
| | | | |
Db 87912 AGCCAAATTAATTAAGCAACTGATTTTGAACAAAGCAACAAACATTAAGTGGGAA 87853
| | | | |

QY 1259 AAGACATCTAGTTAACAAATGATGCTGAGATTTATGGCAACCCATGCTGGAAGATGA 1318
| | | | |
Db 87852 AAGACATCTAGTTAACAAATGATGCTGAGATTTATGGCAACCCATGCTGGAAGATGA 87793
| | | | |
QY 1319 AACTGATCCCTGTCTCTCACTTAATACAAAAATGATACAGATGATCAAGACTTA 1378
| | | | |
Db 87792 AACTGATCCCTGTCTCTCACTTAATACAAAAATGATACAGATGATCAAGACTTA 87734
| | | | |
QY 1379 AATCTGACCTTAAACCAATTAATAATCTAGAGATTAATACAGAAAAATGCTTAGAC 1438
| | | | |
Db 87733 AATCTGACCTTAAACCAATTAATAATCTAGAGATTAATACAGAAAAATGCTTAGAC 87674
| | | | |
QY 1439 ATTCACTTAGGCAAGACTTCAATGCGCAAGAACCCAAAGTAATGCAACAAAAA 1498
| | | | |
Db 87673 ATTCACTTAGGCAAGACTTCAATGCGCAAGAACCCAAAGTAATGCAACAAAAA 87614
| | | | |
QY 1499 ATTAATGATAGGACTTAATTAATCTAATAAGCTTTTGGCAGCAAAAAATCAATAGC 1558
| | | | |
Db 87613 ATTAATGATAGGACTTAATTAATCTAATAAGCTTTTGGCAGCAAAAAATCAATAGC 87558
| | | | |
QY 1559 AGAGCAACAGACACACCCACCGATGAGAGAAATTTCAACAACTAGACATGACTTA 1618
| | | | |
Db 87557 AGAGCAACAGACACACCCACCGATGAGAGAAATTTCAACAACTAGACATGACTTA 87500
| | | | |
QY 1619 GGCATTAATCCGGAATCCCAAGAACTCAACAAATCAGCAGAGAAAGCAAAAT 1678
| | | | |
Db 87499 GGCATTAATCCGGAATCCCAAGAACTCAACAAATCAGCAGAGAAAGCAAAAT 87440
| | | | |
QY 1679 CCCATGAAGAGTGGCTTAAGACATGAATAGCAATTTCTCAAAAGAGATATCAATG 1738
| | | | |
Db 87439 CCCATGAAGAGTGGCTTAAGACATGAATAGCAATTTCTCAAAAGAGATATCAATG 87380
| | | | |
QY 1739 GCCAACAACAGAAAAAATGCTTAATCACTCAATGATTTAGGAAATGTAATCAAC 1798
| | | | |
Db 87379 GCCAACAACAGAAAAAATGCTTAATCACTCAATGATTTAGGAAATGTAATCAAC 87327
| | | | |
QY 1799 TGTATGCGATACCACTTACTCTGCAAGAAATGTCATATTTAAATCTTAATAATTA 1858
| | | | |
Db 87326 TGTATGCGATACCACTTACTCTGCAAGAAATGTCATATTTAAATCTTAATAATTA 87267
| | | | |
QY 1859 TAGATGTTGGTGTCTGTGTGATTAAGAAACATTTTACCTGCTGTGGAAATGTA 1918
| | | | |
Db 87266 TAGATGTTGGTGTCTGTGTGATTAAGAAACATTTTACCTGCTGTGGAAATGTA 87208
| | | | |
QY 1919 ACTTGGCAACCACTATGGAAGAAACAGTGTGAAATTTCTTAAGAACTTAAGTATG 1978
| | | | |
Db 87207 ACTTGGCAACCACTATGGAAGAAACAGTGTGAAATTTCTTAAGAACTTAAGTATG 87148
| | | | |
QY 1979 ACCATTTGATCCAGCAATCCCATTA----- 2003
| | | | |
Db 87147 ACCATTTGATCCAGCAATCCCATTA----- 87088
| | | | |
QY 2004 ----- 2003
| | | | |
Db 87087 CGAAGAAATATCTTGACACACGTTTATAGTACCAATTTGCAATTTGCAATACGTGA 87028
| | | | |
QY 2004 -----AATATGTAATAATTAATTAATTAATTAATTAATTAATTAATTA 2034
| | | | |
Db 87027 ACCAACCAGAAATCCCATCAATCAACAGTGAATTAATTAATTAATTAATTAATTA 86968
| | | | |
QY 2035 TGGAAATCACTCAGCCATTAATAAGATTAATGATGATTCACAGCAATCTAGATG 2094
| | | | |
Db 86967 TGGAAATCACTCAGCCATTAATAAGATTAATGATGATTCACAGCAATCTAGATG 86908
| | | | |
QY 2095 AATTGAGACCTTATTTTAAGTGGGTTAATCTCAGAAATGGAACCAACATCAATATCT 2154
| | | | |
Db 86907 AATTGAGACCTTATTTTAAGTGGGTTAATCTCAGAAATGGAACCAACATCAATATCT 86848
| | | | |
QY 2155 TCTCATCTTAACATGTTGGGTTAAGTGTGAGGACAGAAAGGAT-AGAATGATTAATGA 2213
| | | | |
Db 86847 TCTCATCTTAACATGTTGGGTTAAGTGTGAGGACAGAAAGGAT-AGAATGATTAATGA 86788
| | | | |
QY 2214 ACTCTGGGACT-TGAGGGGAAGATGGAAGAGGCGAGAGATTAAGAACTACAAATG 2272
| | | | |

Db 86787 ACTTGGAGACTCGGGGAAAGGCTGGAGAAAGTAGGTATTAAGACTTACCAATTTG 86728
Qy 2273 GGTACACTGTACACTGTCTGATGGTGCACCAAAATCTCAGAAAT-TACCACTTAA 2331
Db 86727 GGTTCAGTGTATATCTGCCCGGTATGGGTGCACCAAAATCTCAGAAATGAACCTTAA 86668
Qy 2332 GAACCTTATCCATGAGAGCAACACCACTGTTCCCAAAATCCCAATGAATTAATAATAA 2391
Db 86667 GAACCTTACTCATGTACCAAAATACCACTGTTCCCGATACCTATGGAIAAAAAACAA 86608
Qy 2392 TAATATATAA 2401
Db 86607 AAAACAAACA 86598
RESULT 14
AAK68992
ID AAK68992 standard; DNA; 42738 BP.
XX
AC AAK68992;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23804.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytosolic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0214866P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:25:44 ; Search time 692 Seconds

(without alignments)
8856.101 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622

Sequence: 1 tgcattgaagccatgctact.....gtccccaagcccttacc 8622

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTCUTS.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1286.8	14.9	148567	4	US-09-801-876B-3
2	1286.8	14.9	148567	4	US-10-254-869-3
3	1106.8	12.8	32042	3	US-09-245-281-44
4	1106.8	12.8	32042	4	US-09-340-620A-63
5	1106.2	12.8	116592	4	US-09-818-512-3
6	1069.4	12.4	193303	4	US-09-497-855A-37
7	1069.4	12.4	193303	4	US-09-497-855A-44
8	1069	12.4	40328	3	US-08-742-185-102
9	1068.6	12.4	392000	4	US-10-027-983-11
10	1064.6	12.3	168174	4	US-10-071-411A-63
11	1064.6	12.3	168174	4	US-10-071-411A-2
12	1064.2	12.3	99916	4	US-09-816-095-3
13	1058.6	12.3	168174	4	US-10-071-411A-63
14	1058.6	12.3	168174	4	US-10-071-411A-2
15	1054.4	12.2	786431	4	US-09-751-389-3
16	1045.6	12.1	18596	3	US-09-318-448-11
17	1045.6	12.1	18597	4	US-09-962-665-8
18	1045.6	12.1	18597	4	US-09-962-665-8
19	1045.6	12.1	18597	4	US-09-962-665-8
20	984.4	11.4	202001	4	US-09-734-674-3
21	976.8	11.3	10607	1	US-08-078-080-3
22	968.8	11.2	99916	4	US-09-816-095-3
23	949.8	11.0	969	4	US-09-254-227A-4
24	938.6	10.9	969	4	US-09-254-227A-6
25	938	10.9	786431	4	US-09-751-389-3
26	930.6	10.8	168575	4	US-09-426-290-1
27	873.6	10.1	392000	4	US-10-027-983-11

ALIGNMENTS

C	28	863.6	10.0	197496	4	US-09-877-177A-10	Sequence 10, Appl
	29	836.2	9.7	969	4	US-09-254-227A-12	Sequence 12, Appl
	30	833	9.7	969	4	US-09-254-227A-8	Sequence 8, Appl
	31	829.8	9.6	969	4	US-09-254-227A-14	Sequence 14, Appl
	32	826.6	9.6	969	4	US-09-254-227A-10	Sequence 10, Appl
	33	821	9.5	294985	4	US-09-785-381-6	Sequence 6, Appl
	34	802.8	9.3	38844	3	US-09-734-675-3	Sequence 3, Appl
	35	800.8	9.3	29604	3	US-08-781-891-207	Sequence 207, App
	36	800.8	9.3	29604	4	US-09-618-166-207	Sequence 207, App
	37	761.6	8.8	90541	4	US-09-759-359A-3	Sequence 3, Appl
	38	761.6	8.8	90541	4	US-10-207-973-3	Sequence 3, Appl
	39	709.4	8.2	38059	4	US-09-328-925-4	Sequence 4, Appl
	40	706.2	8.2	80246	3	US-09-078-294-4	Sequence 3, Appl
	41	704.6	8.2	80246	3	US-09-078-294-4	Sequence 4, Appl
	42	672.8	7.6	118067	4	US-09-437-855A-32	Sequence 32, Appl
	43	657.2	7.6	15418	4	US-09-783-203-1	Sequence 1, Appl
	44	657.2	7.6	15418	4	US-09-994-427A-1	Sequence 1, Appl
	45	657.2	7.6	15418	4	US-09-244-438-1	Sequence 1, Appl

RESULT 1	US-09-801-876B-3	Sequence 3, Application US/09801876B
Patent No. 6492155	GENERAL INFORMATION:	
APPLICANT: YE, Jane et al	TITLE OR INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC	
TITLE OR INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES		
FILE REFERENCE: CLO01160		
CURRENT FILING DATE: 2001-03-09		
NUMBER OF SEQ ID NOS: 8		
SOFTWARE: FASTSEQ for Windows Version 4.0		
SEQ ID NO 3		
LENGTH: 148567		
TYPE: DNA		
ORGANISM: Human		
FEATURE:		
NAME/KEY: misc_feature		
LOCATION: (1)...(148567)		
OTHER INFORMATION: n = A,T,C or G		
US-09-801-876B-3		
Query Match	14.9%	Score 1286.8; DB 4; Length 148567;
Best Local Similarity	76.3%	Pred. No. 0;
Matches 1972; Conservative	0;	Mismatches 392; Indels 220; Gaps 23;
QY	3	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 60
DB	45746	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 45805
QY	61	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 120
DB	45806	TGCGAGCAATATCATGATGAGACAGATGCTGAATCTTAACAAAATCTAGTAAAC 45865
QY	121	CCATTCACAGCATATCAAGAAATATCCACCATTTCAAGTGGTTTCATCCAGGG 180
DB	45866	TGAATCCAAAGCATATCAAGAAATATCCACCATTTCAAGTGGTTTCATCCAGGG 45925
QY	181	GTGAGG-ATAGGTTAATACATCAAGTCAATTAATGATATCATCAATTAAGAAATT 239
DB	45926	ATGAGGAATGCTTAATCATACATCAATTAATGATATCATCAATTAAGAAATT 45985
QY	240	AAAAAATCAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATTTAA 299
DB	45986	TTTAA-AAAATCAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATTTAA 46044
QY	300	CATTCTTTGATTAACCTTCAGAAATGCAATTAAGAAAGATATA--CCTTAATG 357

Db 46045 CATCCCTTTATGATTAACCCCTCAGCAAAATCAGCATACAAAGGACATAGGCTTAAATG 46104
 Qy 358 TAAATAAAGCCATATATGACGAGCCACAGCAAAACATATATCTGAATGGGAAAAAGTTGA 417
 Db 46105 TAAATAAAGCCATATATGACAAACCCACAGCCAAACATTAACCTGAAC----- 46151
 Qy 418 AAACATGTCCTCCGAGAACTGGAAACAAGCAAGATG-CTACTTTCACCACTTCTATTC 476
 Db 46152 --ACATTCCTCTGAGAACCAAGATGAGACAGATATGCCACTCTCACTGCTCTCTTCA 46209
 Qy 477 ACATAGTAGTGAAGTTTATGACCAAGCAATCAGACAGAAGAAATCAAGGACCC 536
 Db 46210 ATGTAGTACTGGAAGTCTTAGCCAGACATTAAGACAGAAGAAATTAAGTCACTC 46269
 Qy 537 AAATCAATTAAGAGAAAGTCAAACTGTCCCTGTTCACTGATGATATGATTTACTTAG 596
 Db 46270 AAATCAGTAAAGAGAAAGTCAAACTGTCTCTTATGCGATATGATGTTTAACTTG 46329
 Qy 597 AAACCCCTAAAGCTCATCCAGAAAGCTCCTAGAACATGATACATTAATTCAGTAAAGTT 656
 Db 46330 AAACCCCTAAGACCTTTCAGAAAGCTCCTAGAACATGATTAAGAAATTCAGCAAAAGTT 46389
 Qy 657 CAGGATACAACTAAATGTAACAACAAATCAGTAGCACTGCTATACACCAACAGTACCAAG 716
 Db 46390 CCGGATACAAAGATTATGTAACAACAAATCAGTAGCTCTCTATACCAACAGCAACAG 46449
 Qy 717 CTGAAATTAATTAAGAACTCAAACTCTTTCATTAAGCTGTTAA----- 763
 Db 46450 TAGAGAACCAAAATCAAGAACTCAATCCCTTTTACAAATGCTGCAAAAAACAAACAA 46509
 Qy 764 -----AATACTTAAGATATTTCTTACCCAGAG 793
 Db 46510 ACAAGCAAAACAAAAAAACAAAAAAACAAATTAATCTTAGAATATTAACCAAGAG 46569
 Qy 794 GTGAAGGACCTTCTACAA-GGAAAACTACAAACACAGCT---GACATCATAGATGAC 848
 Db 46570 TAGAAAGACCTCTCAACAGGAAAAATTACAAACACCTGGAAGATCATATAGTACAC 46629
 Qy 849 AAACAAGTGGAAACATCCATGCTCATAGGATGGGAGAACTCAATTTGTTG-AAAAATGA 907
 Db 46630 AAACAAGTGGAAACATGCTCCATGCTCATAGGATGAGTAAATCAATGTTTGAATAA 46689
 Qy 908 CCATATGCGCAAAAGCAATCTACAGATTCATGCAATTCACCAAAATTCATCATCAT 967
 Db 46690 CCATATGCGCAAAAGCAATCTACATTAATTCATGCAATTCATCAAAATTCACCATCAT 46749
 Qy 968 TCTTCAAGAACTAGAAAAAAACAATTTCAAAATTCATATGGAACAACCAAAAAA 1027
 Db 46750 TCTTCAAGAAATTAG-AAAAACAATTTCAAAATTCATATGGAAC-----C 46794
 Qy 1028 AAAAAAACCAGCAATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1087
 Db 46795 AAAAAAACCAGCAATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 46854
 Qy 1088 CATTAACCATCTTCAAACTATACAGAGGCTATTAATCAACCAAAACATCATGCACTGAC 1147
 Db 46855 CACTACTGATTTCAATCAATACATTAAGCCCAAGTCAACCAAAACAGCATGATGATGCT 46914
 Qy 1148 ATAAAACTAGGCAATAGCAATGGAAGAAAGAGAAATCCAGAAATTAAGCCAAATA 1207
 Db 46915 AAAAAAATAGGCAATAGCAATGGAAGAAAGAGAAATAGGAACACAGAAATTAAGCTCAATA 46974
 Qy 1208 ATTAATGCAACATGATTTTGAACAAGCAAAACAAATTAAGTGGGAAAAAGACATTC 1267
 Db 46975 CTTAACGCAACATGATTTTGAACAAGCAAAATGAATTAAGTGGAAAAAGACATTC 47034
 Qy 1268 TAGTTAAACAATGATGCTAGATTAATGCAAGCAATGGAAGATGAATGAATGCAATG 1327
 Db 47035 TTTTCAACAATGATGCTAGATTAATGCAAGCAATGGAAGATGAATGAATGCAATG 47093
 Qy 1328 CCTTGTCTCTCACTTAATTAACAAATTTGATACAGATGATCAAGATCTTAATCTGAGA 1387
 Db 47094 CCGTATCTCTCACTTAATTAACAAATTTGATACAGATGATTAAGATCTTAATCTGAGA 47153

Qy 1388 CCTAAACCATTAATTTCTAGAGATTAACATCAGAAAAATGCTCTAGACATTCATTA 1447
 Db 47154 CTTGAACCTTAATTAATTTCTAGAGATTAACATCAGAAAAATGCTCTAGACATTCATTA 47213
 Qy 1448 GCGAAGACCTTATGAGCCAGCAAGAACCAAAAGTAAATGCAACAAAAAATAAATAGA 1507
 Db 47214 AGCAAGGTTTATGATGCAAGAAACCAAAAGCAATGCAATTAATAAAGATTAATTGC 47273
 Qy 1508 TAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1567
 Db 47274 TGGTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47322
 Qy 1568 AGCAACCCAGCGAGTGAAGAAAAATCTTCAAACTAAGCACTGACTAAGCAATTAATA 1627
 Db 47323 AAACGCCCCAGAGTGAAGAAAAATCTTCAAACTAATCTATCTGACAAAGATTAATA 47382
 Qy 1628 TCCGAATCC-ACAAGAACTCAAAACAAATCAGCAAGAAAGCAAAACAAATCCCATGA 1686
 Db 47383 TCCGAATCTCAATGAATCAAGTAAATCAGTAAAGAAA-----AAACAATCTATCA 47438
 Qy 1687 AGAGTGGCTTAAGGACATGAATGACATTTCTCAAAAGAAATTAACAAATGCGCCAA 1746
 Db 47439 AAAGTGGCTTAAGGACATGAATGACAGTTCTCAAAAGAAATTAACAAATGCGCCAA 47498
 Qy 1747 AC--AGAAAAATGCTTAACATCACTAATGATTAAGGAAATGTAATCAACATGTAAT 1804
 Db 47499 ACATAGAAAAATGCTCAACATCACTAATGATTAAGGAAATGCAAAATCAATTAAT 47558
 Qy 1805 GCGATACACCTTACTCTGCAAGAAATGCTCAAT--TTAAAAATCTAAAAATTAATAG 1861
 Db 47559 GTGATTCACCTTACTCTGCAAGAAATGCTTAATTAATTAATTAATTAATTAATTAATTA 47618
 Qy 1862 ATGTTGGTGGTCTGTGTGTGATTAAGAAACATTT-TACATGCTGGTGGGAAATGTAAC 1920
 Db 47619 ATGTTGGTGGTGTGTGTGATTAAGAAACATTTCTACATGCTGGTGGGAAATGTAAC 47678
 Qy 1921 TTGCGCAACCATATGAGAAACAGTGTGGAATTTCTTAAGGAATTAAGATGATGAC 1980
 Db 47679 TAGTACAGCACTATTGAAACAGTGTGGAATTTCTTAAGGAATTAAGATGATGAC 47738
 Qy 1981 CATTTGATCAGCAATGCC----- 1999
 Db 47739 CATTTGATCAGCAATGCCCTTACTGCGTATCTACAGAGAAATTAAGTCAATTAATCA 47798
 Qy 2000 ----- 1999
 Db 47799 AAAAAATTAACAATGATGATTTTACAGACAGAGTTGCAACCAATGCCCCATCA 47858
 Qy 2000 -----ATTAAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2052
 Db 47859 ATCAATGATGATTAAGAAACCTGTGTATATGATTAATTAATTAATTAATTAATTAATTA 47918
 Qy 2053 TAAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2112
 Db 47919 TAAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47978
 Qy 2113 TAAGTGGGTAATCAAGAAATGGAACCAACATCAATGTTCTCACTTAACAATGGGG 2172
 Db 47979 TAAGTGGGTAATCAAGAAATGGAACCAACATCAATGTTCTCACTTAACAATGGGG 48038
 Qy 2173 GCTAAGCTGTAGAGACACAGAGGAT-AGATGATTAATGAACCTGCGGACCTGAGAG 2231
 Db 48039 GCTAAGCTGTAGAGACACAGAGGATTAAGATTAATGAACCTGCGGACCTGAGAG 48098
 Qy 2232 GAAGATGAAGAGAGGCGAGGATTAAGATTAATGAACCTGCGGACCTGAGAG 2291
 Db 48099 GAAGATGAAGAGAGGCGAGGATTAAGATTAATGAACCTGCGGACCTGAGAG 48158
 Qy 2292 AGGTGATGGGTGCAACAAATCTCAGAAATTAATCACTAAGAACTTAATCAATGAGCA 2351
 Db 48159 AGGTGATGGGTGCAACAAATCTCAGAAATTAATCACTAAGAACTTAATCAATGAGCA 48217

Oy 2352 ACACCACTGTGCCCCAAATCCCATGAATAAATTAATTAAATGTTTAAAT 2411
| | | |
Db 48918 A-ACCACCTTTAACCCCATTAAGAAAAATTCAGCACACACTTAGGTTAGTC 48276

Oy 2412 TCAC 2415
||
Db 48277 GGAC 48280

```

RESULT 2
US-10-254-869-3
: Sequence 3, Application US/10254869
: Patent No. 6653117
: GENERAL INFORMATION:
: APPLICANT: YE, Jane et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: FILE REFERENCE: CL001160DIV
: CURRENT APPLICATION NUMBER: US/10/254,869
: CURRENT FILING DATE: 2002-09-26
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 148567
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(148567)
: OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

```

Query Match	14.9%	Score 1286.8;	DB 4;	Length 148567;
Best Local Similarity	76.3%	Pred. No. 0;		
Matches 1972; Conservative	0;	Mismatches 392;	Indels 220;	Gaps 23;

Oy	TATGAAACCAATGTCACTTTAATTCCAAAACCGAGGAAGAATATA--CAAAAAAGAAAAC	60
Dd	TATGAAACCGACATCACCTTAATCCAAAACCATGAAGAACAATACCTTAAAAAGAAAAC	45746
Oy	61 TATGAAACCAAGTACCACATGTAATATATACATGCAGAAATCCCACAACAAATACTAGCTAAC	120
Dd	45806 TGCAAGACCAATATCATTGATGAACAAGATGGTAAATCTTAAACAAANATCTAGCTAAC	45805
Oy	121 CCATCCCAACAGCATATCAAGAAAGATATATCCACCATTTGTCAATGGGTTTCATACCAAGG	180
Dd	45866 TGAATCCAAACGACATATCAAAAAGATATATCCACATGATATAATGGGGTTTCATATACAGG	45925
Oy	181 GTGCAGG-ATAGGTTACATACACAGTCATATAATGTGATACATCAATTAACAGAAAT	239
Dd	45926 ATGCAGGAATGGCTTAATCATACACAAATCATATAATGTGACACACCATTAACAGAAAT	45985
Oy	240 AAAAAACAAAATCATCATGATCATCTTCAATGATGCTGAAAAAGCATTTTGACAAATCTTA	299
Dd	45986 TTTTA-AAAAAATCAATGATCATCTCGTAGGGTCAGAAAAACCATTCACAAAATCCAG	46044
Oy	300 CATTTCTTATGATTAATAACCTTCAGCAAAATGCACATAGAAAAGACAT--CCTTAATG	357
Dd	46045 CATCCTTTATGATTAATAACCTTCAGCAAAATGCACATAGAAAAGACATAGGGCTTTAATG	46104
Oy	358 TAATTAAGGCATATATAGAGGACCCACAGCAACATTATACTGAATGGGGAAAAAGTTGA	417
Dd	46105 TAATTAAGGCATATATAGACAAACCCACAGCAACATTAATACTGAATGGGGAAAAAGTTGA	46151
Oy	418 AAACATTGTCCTGTAGAACTGGAACAAGAGAGATG-CTACTTTGACCACTCTTATTTGA	476
Dd	46152 --AATTGCCCTGTAGAAACGAGATGAGACAGATATGCCACTCTCACTGCTCTCTTTGA	46203
Oy	477 ACATATGATGAGAAAGTTTTAGCCAGACATCAACACAGAAAGAAATACAGGCAACC	536
Dd	46210 ATGTATGATCTGGAAGTCTTACCCAGACCAATTAAGACACAGGAAGAAATTAAGTCAATCT	46268

QY	537	AAATCAATAAAGAGAGTCAAACTGTCCTGGTCACTGATGATATGATTGTAATCCAG	596
Db	46270	AAATCAATAAAGAGAGTCAAACTGTCCTGGTCACTGATGATATGATTGTAATCCAG	46322
QY	597	AAAAACCCATAAGACTCAAGAAAGCTCCATGAACTGATACATATAATTCAGTAAAGTTT	656
Db	46330	AAAAACCCATAAGACTCAAGAAAGCTCCATGAACTGATATAAGAAATTCAGCAAAAGTTT	46380
QY	657	CAGATATCAAACTAAATGTATACAAATCATGTAGCATGTCTATATACCAACAGTGCACAG	716
Db	46390	CCGATATCAAAATTAATGTATACAAATCAGTAGCTCTCTATATACCAACAGCAACCAAG	46440
QY	717	CTGGAATCAAACTCAAGAACTCAAACTTTTACATATAGCTGTAAA-----	763
Db	46450	TAGGAACCAAACTCAAGAACTCAAACTTTTACATATAGCTGTAAA-----	46500
QY	764	-----AATACTTAAGATATATTTCTTCCCAAGAG	793
Db	46510	ACAAAGCAAAACAAAAAAACAAAAAAACAAATATCTTAGAATATATCTTAACCAAGAG	46560
QY	794	GTGAAGACCTCTACAA-GGAAAACTACAAACAAGCT----GACATCATAGATGACAC	848
Db	46570	TAGAAAGACCTCTACAAAGGAAAAATTACAAACACTGCTGGAGAAATCATATATGACAC	46620
QY	849	AAACAAGTGGAAACATCCATGCTCATGTGATGGATGAAATCAATATTGTC-AAAAATGA	907
Db	46630	AAACAAGTGGAAACATGCTCCATGCTCATGTGATGGATGAAATCAATATTGTC-AAAAATGA	46680
QY	908	CCATATGCTCCAAAGCACTACAAAGTTCATGCAATCCACCAAAATATCATCATCAT	967
Db	46690	CCATATGCTCCAAAGCACTATATTAATTCATGCAATTTCCATCAAAATATCATCATCAT	46740
QY	968	TCTTCAAGAACTGAAAAAAACAAATCTTAAATTCATATGGAACAAACAAACAAAAAA	1027
Db	46750	TCTTCAAGAAATG-AAAAACAAATCTTAAATTCATATGGAAC-----C	46797
QY	1028	AAAAAAACCCGATAGCCAAAGCAAGCTTAGCAAAAGAACAAATCTGAGGCATCA	1087
Db	46795	AAAAAAACCCGATATAGCCAAAGCAAGCTTAGCAAAAGAACAAATCTGAGGCATCA	46850
QY	1088	CATTAACCATTTGAAACTATATCAAGAGTATATACCAAAACATCATGAGCACTGAC	1147
Db	46855	CATTAACCATTTGAAACTATATCAAGAGTATATACCAAAACATCATGAGCACTGAC	46910
QY	1148	ATTAACACTAGGCATATAGCAACATGSAAGAAAGAGAGATCCAGAAATTAAGCCATATA	1207
Db	46915	ACAAATATAGGCATATAGCCAAATGSAAGAGAAATGSAAGACACAGAAATTAAGCCATATA	46970
QY	1208	ATTAATAGCAACTGATTTTGAACAAGCAAAACAAATCATAAAGTGGGAAAGACATTC	1267
Db	46975	CTTAACAGCAACTGATTTTGAACAAGCAAAACATTAAGTGGGAAAGACACATTC	47030
QY	1268	TAGTTAACAAATGCTGCTGAGATTAATTGGCAAGCCACATGTGGAAAGATGAATCGATC	1327
Db	47035	TTTTCACAAATGCTGCTGAGATTAATTGGCAAGCCACATGTGGAAAGATGAATCGATC	47090
QY	1328	CCTTGCTCTCACTTAATACAAATTAATGATACAGATGATCAAGACTTAATCTGAGA	1387
Db	47094	CCTTGCTCTCACTTAATACAAATTAATGATACAGATGATGAATTAAGACTTAATCTGAGA	47150
QY	1388	CCTAAACCATTAATAATTTCTAGAAAGATACATCAGAAAAATGCTTCTAGACATTCATTA	1447
Db	47154	CCTGAACCTAATAATAATTTCTAGAAAGATACATCAGAAAAATGCTTCTAGACATTCATTA	47210
QY	1448	GGCAAAAGCTCATGAGCCAGAAACCCCAAAAGTAAATGTCAACAAAAACAAATTAATGTGA	1507
Db	47214	AGCAAGGATTCATGAGCCAGAAACCCCAAAAGTAAATGTCAACAAAAACAAATTAATGTGC	47270
QY	1508	TAGAGCTTAATTAATTAAGCTTTTGGCAGCAAAAAACATCATTTAGCAGAGCAAAAC	1567
Db	47274	TGATATCTTAATTAATTAAGAGCTTTTGGCAGCAAAACGAG-----AGTCAGC	47320
QY	1568	AGACAAACCAACGAGTGAAGAAAAATCTTCAAACTAAGATCTGACTAAGACTAATA	1627


```

QY 358 TATATAAGCCATATATGACGACCCACAGCAAACTTATATGTAATGGGAAAGTGA 417
DB 14292 TATATAGAGCTATCTATGACAAACCCAGCCAAATCATCTGTAATGGCAAAAAGTGG 14233
QY 418 AAAACATGTCCTGAGAACTGGAACAAAGCAAGATG-CTACTTTCACCACTTCTATTC 476
DB 14232 AAGCATTCCTTGGAAAACCTGGCAAGACAGGAGTCCCTCTCTCCGCTCTATATCA 14173
QY 477 ACATGTAGTGAAGTTTATAGCCAGAGCATGACAGAGAAAGAAATCAAGGCAACC 536
DB 14172 ACATGTAGTGAAGTTTATAGCCAGAGCATGAGGAGAAAGAAATCAAGGATTC 14113
QY 537 AAATCAATTAAGAGAGATCAAACTGTCCTGTTCACTGATGATGATGATTTACCTAG 596
DB 14112 AATTGAGAAAGAGAGATCAAAATGTCCTGTTTCAAGACGACATGATGTTTATCTAG 14053
QY 597 AAAACCTTAAGAATCATCCAGAAAGCTCTAGAACTGATATCAATTAATTCAGTAAAGTTT 656
DB 14052 AAAACCCCATGCTCAGCCCAAAATCTCTTAAGCTGATTAAGCAACTTCAGCAAGATCT 13993
QY 657 CAGGATTCAGAACTTAATGTCACAAATCATGATGACATGCTATACCCAAAGTACCAAG 716
DB 13992 CAGGATTCAGAAATCAATGTCAGAAATCAAGACATCTTATACACCAAGAGAGCAAA 13933
QY 717 CTGAGATCAAAATCAAGAACTCAACACTTTCATTAAGCTGTAAA-----AAAATACT 770
DB 13932 CAGAGAGCCAAATCATGAGGAGAACTCCATTCATTAATGCTTCAAGAGAAATTAATATCC 13873
QY 771 TAAAGATATTTCTTAAACCAAGAGGTGAAGAGCTCTACAGAGAAACTACAAACACAGC 830
DB 13872 TGAAGATCAAACTTAAGAGGATGTGAAGAGCTCTTCAAGAGAACTACAAACCACTGC 13813
QY 831 T-----GACATCTAGATGACACAAAGAGTGAAGAAACATCCATGCTCATGATGGGTA 886
DB 13812 TGAAGAAATTAAGAGAGACAAACAAATGAGAAACATTCATGCTCATGAGGTAGAA 13753
QY 887 GAATCAATATTTGAAATAGACCATATTTGCCAAAGCAATCTACAGTTCAATGCAATTC 946
DB 13752 GAATCAATATCTGTAATAGCCATATCTGCCAAAGTAATTTACAGTTCAATGCAATTC 13693
QY 947 CCACCAAAATATCATCATCTTCTTCAAGAACTGAGAAACCAATTTCTTAAATTCATA 1006
DB 13692 CCATCAAGCTACCAATGACCTTCTTCAAGAA--TTGGAATAAACTACTTTAAAGTTCATA 13634
QY 1007 TGAACCAAGCAACCAAAAAAAGAAAAAAGCCGCAATGCAAGCAAGCAAGCTTAAGCAAA 1066
DB 13633 TGGAAAC-----CAAAAAAGACCCGCAATGCAAGTCAATCTTAAGCCAAA 13588
QY 1067 AGAACCAATCTGAGGCAATCACTTACCCATCTTCAAACTATCTACAGAGCTATATCA 1126
DB 13587 AGAACCAAGCTGAGGCAATCACTTACCCATCTTCAAACTATCTACAGAGCTATATCA 13528
QY 1127 CCAAAACATCATGCACTGACATTAACCTAGGCAATGACCAATGGAAGAAAGAGAGA 1186
DB 13527 CCAAAACACATGCTGCTGTCACCAAAAGAGATATGATCATGAGAAAGAGAGAGC 13468
QY 1187 ATCCGAAATTAAGCAAAATTAATATAGCAAGCAATGTTTGAAGCAAGCAAGCAAGAA 1246
DB 13467 CCTCGAAATTAAGCGGCAATATCTACATCTGATCTTGAAGCAAGCTGAGAGAAACA 13408
QY 1247 TAAAGTGGGAAAGACATTTCTAGTTTAAACAAATGCTGAGATTAATGCAAGCCACAT 1306
DB 13407 AGCAATGGGAAAGAGATTCCTTATTTAATAGTGTCTGGAAGAACTGGCTAGCCATAT 13348
QY 1307 GTGGAAGAAATGAATGAGATCCTTGTCTCTCACTTAATACAAAAATTTGATACAGATG 1366
DB 13347 GTAGAAAGCTGAAGATGATCCCTTC-CTTACACCTTATCAAAAAATCAATTCAGATG 13289
QY 1367 ATCAAGAGCTTAATCTGAGAACTTAAGCAATTAATTAATTAAGAAATTAATCAAGAAA 1426
DB 13288 ATTAAGAGATTTAAACGTTAAACCTTAAGCAATTAAGAACTTGAAGAAAGCACTTGA 13229
QY 1427 ATGCTTCTAGACATTCATTAAGCAAGAACTTCAATGAGCAAGAACCAAAAGTAAATGA 1486

```

```

DB 13228 ACATTCAGAGACATAGGCGGTGGCAAGAGACTTCATGTCAAAACACCAAAAGCAATGCA 13169
QY 1487 ACMAAAACAAAATTAATATGATAGCACTTAATTAATTAATTAAGCTTTTGGCGCAAAA 1546
DB 13168 ACMAAAAGCAAAATTAAGCAAAATGAGATCTTAATTAATTAAGAGCTTCTGCAAGCAAAA 13109
QY 1547 ACATCATTAAGAGAGCAAGCAAGCAACCCAGAGTGAAGAAATCTTCAAAACTPA 1606
DB 13108 GAATCTACATCAAGAGTGAACAGGCACTTACAAATGAGAGAAATTTGGCAACCTAC 13049
QY 1607 GCATCTGACAAAGCACTAATATCCGAAATCCAAAGAACTTCAACAAATGCAAGAAAG 1666
DB 13048 TCATCTGACAAAGGCTTAATATCCAAATCTCAATGAACTCAACAAATTAATTAAGAAA 12989
QY 1667 AAGCAAAACAAATCCCATGAAGAGTGGGTAAAGACATGAATAGCAATTTCTCAAAAGAA 1726
DB 12988 AAAACAAACAAACCCCATCAAAAGTGGGGAAGCATGAACGACCTTCTCAAAAGAA 12929
QY 1727 GATATCAAAATGCGCAACA--ACAGAAATAATGCTTAACATCACTAATGATTAAGGAA 1784
DB 12928 GACATTTATGACGCAAAACCAATGAGATATCATCTACACCAAGTTAAGATGCAATTA 12869
QY 1785 ATGTAATCAACACTGTAAATGCAATACCACTTACTCTGCAAGATGTCATAATTTAA 1844
DB 12868 ATGCAATCAAAACCACTAATGATATCATCTACACCAAGTTAAGATGCAATTA 12809
QY 1845 AATCTAAATTAATGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1904
DB 12808 AAGTCAGAAACAGAGGCTGTGAGAGATGAGAGAAATGAAGCACTTTACACTGT 12749
QY 1905 TGGTGGAAATGTAAATCTTCCGCAACCACTATGAAAGCAAGTGTGAAATTTCTTAAGAA 1964
DB 12748 TGGTGGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12689
QY 1965 CTAAAGTATGATGACATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2024
DB 12688 CTGAAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12630
QY 2025 TTTATATCATG 2036
DB 12629 GTATTAATCATG 12618

RESULT 5
US-09-818-512-3/C
; Sequence 3, Application US/09818512
; Patent No. 6537780
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001192
; CURRENT APPLICATION NUMBER: US/09/818,512
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

Query Match 12.8%; Score 1106.2; DB 4; Length 116592;
Best Local Similarity 74.6%; Pred. No. 5.4e-271;
Matches 1531; Conservative 0; Mismatches 488; Indels 32; Gaps 10;

QY 1 TGTATGAGCAATGTCACTTTAATACAAAACAGAAAGATATA--CAAAAAAGAA 58

```


Db 51038 TTATATGAGGCGACATCATTTCTGATACCAAGCCAGGACGACACACCAAAAAAGGA 50979
 Qy 59 ACTATAGACCAAGTACCACTGATGAATATACATGACGAATCCCAACAAATACTACTA 118
 Db 50978 ATTTTACCAATATCTCTGATGAACATGATGCAAAAAATCTCAATAAATATCTGGCA 50919
 Qy 119 ACCCAATCCACACGATATCAAGAAATATCCACATTTGCAAGTGGGTTTCAATCCAG 178
 Db 50918 ACTGAATCCAGACGACATCAAAAAAGCTATCCACATGATCAGGTGGCTTCAATCCCTG 50859
 Qy 179 GGGTGC-AGGATAGTTTAACTACACAAAGTAAATGATATCATCATCAATAACAGAA 237
 Db 50858 GGATGCAAGGTGGTTCAATATACCAAAATCAATTAATGTAATCCAGCATATTAACAGAG 50799
 Qy 238 TTAATAACAAAAATCACATGATCATCTCAATAGATGCTGAATAAAAGCATTTGACAAATCT 297
 Db 50798 CCAAGACAAAAACACATGATTTATCTCAATAGATGCAAAAAAGCCCTTGAACAAATTC 50739
 Qy 298 AACATTTCTTATGATTAATAAATCTTACGCAAAATCGACATAGAAAGACATACCTTAATG 357
 Db 50738 AACCACTTCATCTTAAAACTCTCAATAAATAGGATTTGATGGGACGTATCTCAAAA 50679
 Qy 358 TAAATAAAGGCATATATAGGAGGACCCACAGCAACATTATATCTGAATGGGAAAGTTGA 417
 Db 50678 TAAATAAGACATATATATGACAAACCCACAGCAATATCATATGAATGGGCAAAACCTGG 50619
 Qy 418 AACATTTCTCTGAGATGGAACAAAGACAGATG-CTACTTTCACCACTTCTATTTCA 476
 Db 50618 AAGATTTCTCTTGAATAATGCGACAGACAGGATGCTCTCTCACTGCTCTATTTCA 50559
 Qy 477 ACATATGATGGAAGTTTAAAGCCAGAGCAATCAGACAGAAAGAAATCAAGGCAACC 536
 Db 50558 ACATAGGTTGGAAGTTCTGGCCAGGCAATCAGGACAGAAAGAAATTAAGGATATTC 50499
 Qy 537 AAATCAATTAAGGAGAGTCAAACTGTCCCTGTCATGATGATATGTTATATCTAG 596
 Db 50498 AATTAGAAGAGAGAGTCAAAATGTTCTGTTGACAGACATGATGTTATATCTAG 50439
 Qy 597 AAAACCTTAAGACTCATCTCCAGAAAGCTCTAGATGATATCATAAATTCAGTAAAGTT 656
 Db 50438 AAAACCCATGCTCTGAGCCCAAAATCTCTTAAGCTGATAGCACTTCAGCAAAAGCT 50379
 Qy 657 CAGATTCAAACTAATATGATACAAATCATAGTACATGCTATACCAACAGTACCAAG 716
 Db 50378 CAGATTCAAAAATCAATGTAACAAAAATCAACAGCATTTCTATACCAACAGACAAA 50319
 Qy 717 CTGAGATCAAAATCAAGAACTCAAACTTTAACAATAGCTGTAAA-----AAATATCT 770
 Db 50318 CAGAGACCCAAATCATAGGTGAATCTCCATTCACAATTTGCTTCAAGAGAAATTAATATCC 50259
 Qy 771 TAAAGATATTTCTTACCAAGAGGTGAAGACCTCTACAGAGAAACTACAAACACAGC 830
 Db 50258 TAGGAATCCAACTTACAGAGATGTAAGGACCTCTTCAAGAGAACTTCAACAACTGC 50199
 Qy 831 T----GACATCATATGATGACAAACAGTGAACACATCCATGCTCATGATGGGTA 886
 Db 50198 TCAAGGAAATTAAGAGACACAAACAAATGAGAAACATTCATGCTCAGGGTAGGAA 50139
 Qy 887 GAATCAATATTTGGAAGAAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATTC 946
 Db 50138 GAATCAATATTTGGAAGAAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATTC 50079
 Qy 947 CCAACCAATATATCATCATTTCTTCAAGAACTAGAAAAAAACATTTCTAAATTTGATA 1006
 Db 50078 CCAATCAAGCTACCAATGATCTTTCTTCAAGAA-TTGGAAGAAATCTATTTAAAGTTGATA 50020
 Qy 1007 TGAACCAACACCAAAAAAAACCCGCAATAGCCAAAGCAAGACTTAGCAAAA 1066
 Db 50019 TGGAAAC-----CAAAAAAGAGCCCAATTTGCAAGTCAATCTTAAGCAAAA 49974
 Qy 1067 AGAACAATCTGAGGATGACATTAACCATCTTCAACTTACTTACAGGCTTATATCA 1126

Db 49973 AGAAACAAAGCTGAGGACATCACTACTGACTTCAAACTTACTACAGGCTACAGTA 49914
 Qy 1127 CCAAAACATCATGCACTGACATTAATACTAGGACATAGACCAATGGAAGAAAGAGA 1186
 Db 49913 CCAAAACAGATGCTATCTGGTACCAAAACAGAGTATATGATCATATGGAACAGAACAGC 49854
 Qy 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAAAACAAA 1246
 Db 49853 CCTCAGAAATTAATGCCCCATATCTACAACTATCTGATCTTTGAACAAACCTGAGAAAA 49794
 Qy 1247 TAAAGTGGGAAAGACATTTAGTTAACAATAGTGTCTGAGATTAATTTGGCAAGCCAT 1306
 Db 49793 AGCAATGGGAAAGATTCCTCTATTTAATTAATGCTCTGGAAACCTGGCTACCATAT 49734
 Qy 1307 GTGGAAGATGAATCTGATCCCTTGTCTCTCACTTAATCAAAAAATTGATACAAGATGG 1366
 Db 49733 GTGAAAGCTGAATCTGATCCCTTC-CTTAACCTTATACAAAAATCAATTAAGATGG 49675
 Qy 1367 ATCAAGACTTAAATCTGAGACTTAAACCAATTAATAATTTAGAAATTAACATGAAAA 1426
 Db 49674 ATTAAGATTTAAACGTTAAACCTTAAACCATTAACAAACCTAGAAAGAAACCTAGGCA 49615
 Qy 1427 ATGCTTACAGCACTTACCTTAGGCAAGACTTCAATGCAAGAACCCAAAGTAAATGCA 1486
 Db 49614 ACATTCAGACATAGGCGTGGGCAAGACTTCAATGCAAGAACCAAAAGCAATGCA 49555
 Qy 1487 ACAAAAAAATAATTAATAGTATGAGCTTAATTAATACTAAAGCTTTTGGACGAAAA 1546
 Db 49554 ACAAAACAAATTTGACAAATGAGATCTAATTAATTAATAAGCTTCTGACAGCAAAA 49495
 Qy 1547 ACAATCTTATGACAGCAACACAGAACCAACCGAGTGAAGAAATCTTCAAACTTA 1606
 Db 49494 GAAACTCAATCAGAGTGAACAGGCACTTACAAATGGAAGAAATTTTGCATCTTAC 49435
 Qy 1607 GCATCTGACTTAAGGACTTAATATCGGAATCCAAAGGACCTCAACAAATCAGCAAGAG 1666
 Db 49434 TCAATCAGAAAGGCTTAATATCCAGATCTACATATTACTCAACAAATTTACAGAAA 49375
 Qy 1667 AAAGCAACATCCCATGAAGAGTGGCTTAAGGACATGATGACAAATCTCAAAAGAA 1726
 Db 49374 AAAACAAACACCCATCAAAAGTGGCGAAGACATGAAACAGACCTTCTCAAAAGAA 49315
 Qy 1727 GATATCAAAATGCGCAAAA-CAGAAAAATGCTTAATCACTCTATGATTTGGAAA 1785
 Db 49314 GACATTTATCAGCAAAAAACCATGAAGAAATGCTCATGACTGCGCACTGAGAGAA 49255
 Qy 1786 TGTAAATCAACACTGTAATGATGATACACCTTACTCTGCAAGAAATGTCATATTTAAA 1845
 Db 49254 TGCATTAACAAACCATATGATGATATCTCACACAGGTGAGATGCAATCTAABAA 49195
 Qy 1846 AATCTAAATAATATAGATGTTGGTGTCTGTGTGTAAGAAACCTTTTACACTGCT 1905
 Db 49194 AGTAGAAGAACACAGGTGTCTGAGAGATGCGGAGAAATAGAAACCTTTTACACTGTT 49135
 Qy 1906 GTTGGATGTAACTTGGCGAACCATATGAAAAACAGTGTGAAATTTCTTAAGAAC 1965
 Db 49134 GTTGGACTGTAACTGTTCATCAACTATGGAAGTCAAGTGTGCGATCTCTCAGGGAATC 49075
 Qy 1966 TAAAGTATGATGACCAATTTGATCCAGCAATCCATTAATATGATTAATATATATTT 2025
 Db 49074 TAGAATGAAATATCAATTTGACCAAGCAATCCAT-ACTGGGTAATATCCCAAAATGAG 49016
 Qy 2026 TATATACCATG 2036
 Db 49015 TATTAATCATG 49005

RESULT 6
 US-09-497-855A-37/c
 ; Sequence 37, Application US/09497855A
 ; Patent No. 6605432
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Tim

```

; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UWO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match      12.4%; Score 1069.4; DB 4; Length 193303;
Best Local Similarity 74.1%; Pred. No. 1.9e-261;
Matches 1497; Conservative 0; Mismatches 486; Indels 36; Gaps 10;

QY 1 TGTATGAGCCCAATGCTACTTTATATACCAAAACGAGAAAGATAT-CAAAAAGAA 58
DB 93923 TTTATGAGCCGACATCTCTATATACAAAGCGTGGCAGAGACACAAACCAAAAGAGA 93864
QY 59 ACTATGAGCCAGTACCACTGATATATATACATGACAGAAATCCCAACAAATCTAGCTA 118
DB 93863 ATTTGAGCAAAATCTCTGATGAAATGATGACAAAAATCCCAATAAATCTGCA 93804
QY 119 ACCCAATCCCAAGCATATCAAGAAATATCCACATTTGTCAAGTGGTTTATACAG 178
DB 93803 ACCCAATCCGACGACATCAAAAGCTTATTCACATGATCAAGTGGCTTATCCGTG 93744
QY 179 GGGTGC-AGGATAGGTTAACTATACACAGTCAATTAATGTGATACATCACTAAACGAA 237
DB 93743 GGATGCAAGGCTGTTCAATATATAGCAAAATCAATTAATGTATCCAGCATATTAACGAA 93684
QY 238 TTTAAACCAAAATCATGATCATCTCATATAGATGTGAAAGCAATTTGACAAATCT 297
DB 93683 CCAAGACAGAAACCAATGATATATCTCAATGATGAGAAAGCTTTTTCACAAATTC 93624
QY 298 AACATTTCTTTATGATTTAAACCTTCAGCAAAATTCACATAGAAAGACATACCTTATG 357
DB 93623 AACCAACCTTCAAGTACAAACCTCTCAATTAATTAAGTATGATGGAACGATCTCAAA 93564
QY 358 TAATAAAGCCATATATGACGACACCAACGAAACATTATATCTGATGAGGAAAGTTGA 417
DB 93563 TAATAAGACTATCTATGACAAACCCACACCAATATCATATGAAAGGCAAAACCTGG 93504
QY 418 AAAACATTTGCCGAGAACTGGAACGAAACAGAGAGT-CTACTTTCACCACTTCTATTC 476
DB 93503 AAGCATTTCCCTTTGAAACCTGGCAAGACAGAGGATCCCTCTCTCACTCTATTC 93444
QY 477 ACATATAGTGAAGTTTATGACGAGACATCAGACAGAAAGAAATCAAGGGCACCC 536
DB 93443 ACATATGTTGGAAGTTCTGGCAGGGCAATTAGGACAGAGAAATTAAGGTATTC 93384
QY 537 AATCAATAAAGAGAAAGTCAAACTGTCTCTGTTCACTGATGATATGATTTATCTAG 596
DB 93383 AATTAGGAAAGAGAAAGTCAAAATTTGCTGTTGAGATGACATATGTTATTAACCTAG 93324
QY 597 AAAACCTTAAGATCATCCAGAAAGCTCCTAGAACGATGATCAATTAATTCAGTAAATTT 656
DB 93323 AAAACCCATGCTCAGCCCAAAATCTACTTAAGCTGATGAGCACTTCAGCAAACTCT 93264
QY 657 CAGGATACAAACTAAATGTACACAAATCAGTACGTATATACACCAAGTACCAAG 716
DB 93263 CAGGATATTAATATCATATGTAACAAATCAAGCATTCATATACCA-----TAACAA 93208
QY 717 CTGAGATCAAAATCAAGAACTCAAAACCTTTTACATAGCTGTAA-----AAATACT 770
DB 93207 CAGAGAGCCAAATCATGATGAACTCCCATTCACAAATGCTTCAAAAGGAAATTAATATCC 93148
QY 771 TAAAGATATTTCTTACCAAGAGGTGAAGACTCTTACAGAGAAATCTACAAACACAGC 830

```

```

DB 93147 TAGGAATCAACTTACAGAGGACATGAAGGGCTCTTCAAGAGAACTACAAACCACTGC 93088
QY 831 T-----GACATCATGATGACACAAACAGTGAACACATCCCATGCTCAGATGGTA 886
DB 93087 TCAGAGAAATTAAGAAAGATTCAAAGAAATGAAGACTATCCATGCTCAGGCTAGGAA 93028
QY 887 GAATCAATATTTGAAATGACCATATGCGCAAAAGCAATCTACAGTTCAATGCAATTC 946
DB 93027 GAATCAATATCTGTAATTTGGCCATATCTGCCAGTAAATTAATGATTCAGTCCATTC 92968
QY 947 CCACCAATATATCATCATCTATTCTTACAGAACTAGAAAAAACAATTTAAATTCATA 1006
DB 92967 CCATCAAGCTACCAATGACTTTCTTACAGAAATGAAAAAATCACTT-TAAAGTCAATA 92909
QY 1007 TGGAAACACACCAAAAAAAAAAAAAAAAAAAAAAACCCTGATGCCAAAGCACTTGGCAAA 1066
DB 92908 TGGAAAC-----AAAAAAAAAGCTGATCACAAGTCAATCTTAAGCCATA 92863
QY 1067 AGAACAAATCTGAGGCATCACTTACCATCTTCAACTATATCTACAGGCTATTAATCA 1126
DB 92862 AGAACAAAGCTGAGGCATCAAGCTTACCTGACTTCAACTATATCTACAGGCTACAGTA 92803
QY 1127 CCAAAACATCATGCACTGACATTAATACTAGGACATAGACCAATGAAAAAGAGAGA 1186
DB 92802 CCAAAACGATGATGATCTGATCAAAACAGAGATATGATCAATGAAACAGAACAGC 92743
QY 1187 ATCCGAATTAAGCCAAATATTTATGACCACTGATTTTGAACAAACCAAAACA 1246
DB 92742 CCTCAGAAATTAATGCTGCTATCTCAACTATCTATCTTGAACAACTGAGAAACA 92683
QY 1247 TAAAGTGGGAAAGACATTTAGTTAACTAAATGAGTGTGAGATTAATGGCAAGCAT 1306
DB 92682 AGCAATGGGAAAGGATTCCTATTTATTAATGATGCTGGAAATCTGCTACGATAT 92623
QY 1307 GTGGAAGATGAAATGATGATCCCTTGTCTCTCACTTAATCAAAATTTGATACAGATG 1366
DB 92622 GTAGAAAGCTGAAATGATGATCCCTT-CCTTACACTTATACAAATTAATTAAGATG 92564
QY 1367 ATCAAAAGCTTAATTTGAGACTTAACCAATTAATTTCTGAGATTAATCAAGAA 1426
DB 92563 ATTAAGACTTAATTTGAGACTTAACCAATTAATTTCTGAGATTAATCAAGAA 92504
QY 1427 ATGCTTCTGACATTCATCTTAGCAAGACTCATGSCCAAGAACCCAAAGTAATGCA 1486
DB 92503 ACCATTCAGACATGAGCATGAGCAAGACTTCAATCTTAAACACCAAAAGCAATGCA 92444
QY 1487 ACNAAACAAATTAATGATAGACTTAATTAATTAATTAAGCTTTTGGCAGCAAA 1546
DB 92443 ACNAAAGCCAAATTAAGCAATTCGATTAATTAATTAAGCTTTCGACACAAAA 92384
QY 1547 ACATCATTAAGCAGACAAACGACACCCAGAGTGAAGAAATCTTCAAACTTA 1606
DB 92383 GAAACTACATCAGACTGAAACGAGCAACCTTAACAAATGAGAAATTTTCCCAACTAC 92324
QY 1607 GCATGCACTAAGGACTAATATTCGGGAATCCACAGGAATCAACAAATCAGCAAGAG 1666
DB 92323 TCATTCGACAAAGGCTGATATCCAAATCTCAATGAACTCCAAATTTTCAAGAA 92264
QY 1667 AAAGCAACATTCATTCAGAAAGTGGGCTTAAGACATGATAGCAATTCCTCAAGAA 1726
DB 92263 AAAGCAACATTCATTCAGAAAGTGGGCAAGATGACATGACATCTTCAAGAA 92204
QY 1727 GATATACAAATGGCACA--ACAGAAAAAATGTTAACTCACTAATGATTGGAA 1784
DB 92203 GACATTTATGACACCAAAACACATGAAATAATGCTCACATCTGCGCCATGAGAA 92144
QY 1785 ATGTAATCAACCTGATATGATACCACTTACCTTCCGCAAGATGCTAATTTTA 1844
DB 92143 ATGCAATCAAAACCAATGATATCACTCACACGTTTGAATGCAATCAATTA 92084
QY 1845 AATCTTAAATTAATGATGTTGTGGTCTGTGATTAAGAAACATTTTACACTGC 1904

```

Db 92083 AAGTCAGAAACCAAGTGTCTGAGAGATGTGAGAAATAGAACCTTTACACTGT 92024
 QY 1905 TGGTGGAAATGTAACTTGGCCAAACCACTATGAGAAAAGTGTGAAATTTCTTAAGGA 1964
 Db 92023 TGGTGGAGCTGTAAACTAGTTCAACCTTGTGGAAGTCAGTGTGGCCATTCCTCAGGAT 91864
 QY 1965 CTAAAGATAGTCCAGCCCTTTGATCCAGCAATCCCATTA 2003
 Db 91963 CTAGAACTAGAAATACATTTGACCCAGCCATCCCATTA 91925

RESULT 7

US-09-497-855A-44/C
 ; Sequence 44, Application US/09497855A
 ; Patent No. 6605432
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Tim
 ; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
 ; FILE REFERENCE: UMO1523
 ; CURRENT APPLICATION NUMBER: US/09/497, 855A
 ; CURRENT FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: 60/120,592
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: 60/118,760
 ; PRIOR FILING DATE: 1999-02-05
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 193303
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens;
 ; US-09-497-855A-44

Query Match 12.4%; Score 1069.4; DB 4; Length 193303;
 Best Local Similarity 74.1%; Pred. No. 1.9e-261;
 Matches 1497; Conservative 0; Mismatches 486; Indels 36; Gaps 10;

QY 1 TGTATGAAGCCATATGTCACTTAATACCAAAACAGAAAGATATA--CAAAAAGAA 58
 Db 93923 TTTATGGGCGCAGATCATCTTAATACCAAGCCTGGCAGAGACAAACAAAAGGA 93864
 QY 59 ACTATAGACGATGCACTGATGATATATACATGAGAAATCCCAACAAATACTAGTA 118
 Db 93863 ATTTGGACCAATATCTTGATGAAACATGATGAGAAAATCTCAATAAATACGCGAA 93804
 QY 119 ACCCAATCCACAGCATATCAAGAAATATCCACATTTCAAGTGGTTTCATACAG 178
 Db 93803 ACCAAATCCAGCAGCATCAAAAAGTTATTCACATGATCAAGTGGCTTCATCCGTG 93744
 QY 179 GGGTGC-AGGATAGATTAACTATACACAGATCAATTAATGTGATACATCACTAAACAGAA 237
 Db 93743 GGAATGCAAGGCTGTTCATATATGCGAAATCAATTAATGTATACAGATTAATTAACGAA 93684
 QY 238 TTAATAACAAATACATGATCATCTCAATAGATGCGAAAAGCATTTGACAAATCT 297
 Db 93683 CCAAGACGAAACCACTGATTAATCTCAATAGATGCGAAAAGCTTTTCACAAAATTC 93624
 QY 298 AACATTTCTTATGATTTAAACCTTCAAGCAAAATGCAATAGAAAGCATACCTTAATG 357
 Db 93623 AACAACTTCAAGCTCAACCTCTCAATATTAAGTATGATGAGCGATATCTCAAAA 93564
 QY 358 TAATTAAGGCATATATAGACGACCCACAGCAAACTATTAATGATGGGAAAAAGTTGA 417
 Db 93563 TAATTAAGGCATATATAGACCAACCCACAGCAATATCATCTGAATGGGCAAAAAGCTGG 93504
 QY 418 AACATTTCTTCAAGCTGGAACAGCAAGATG-CTACTTTCACCACTTCTATTTCA 476
 Db 93503 AAGCATTTCTTTAAACCTGCAACAGACGAGATGCCCTCTCTCAACCACTCTATTTCA 93444
 QY 477 ACATATAGTGAAGTTTATGACAGAGCAATCAGACAGAAAGAAATCAAGGCAACC 536
 Db 93443 ACATATAGTGAAGTTCTGCGCAGGCGAATTAAGCAGAGAAAGAAATTAAGGTAATTC 93384

QY 537 AAATCAATTAAGAGAACTCAAATGTCCTGTTCACTGATGATATGATTTGTATACCTAG 596
 Db 93383 AATTAGGAAAGAGAAAGTCAAAATTTGCTCCGTGTTGACAGATGATGATTTATATACCTAG 93324
 QY 597 AAAACCTTAAGATCTCATCCAGAAAGCTCCAGAACTGATATCATTAATTCAGTAAGTTT 656
 Db 93323 AAABCCATGCTCTACGCCCAAAATCTTAAAGTATAGCACTTCACAGAAAGTCT 93264
 QY 657 CAGGATCAAACTTAATGTACAAATCAGTAGCACTGATATACCAACAGTACCAAG 716
 Db 93263 CAGGATATTAATTCATGATCAAAAATACAAAGATTCATATACCA--TACGAA 93208
 QY 717 CTGGAATCAATCAAGAACTCAACACTTTTACAAATGCTGTAA-----AAATGCT 770
 Db 93207 CAGAGACCAATCATGATGAACTCCCATTCACAATTTGCTTCAAGAGAAATTAATATCC 93148
 QY 771 TAAGATATTCTTCCCAAGAGGTGAGAGACCTCTCAAGAGAAACATCAAAACAGC 830
 Db 93147 TAGGAATCCAACTTACAAAGGACATGAAAGGCTCTTCAGAGAACTTCAAAACCACTGC 93088
 QY 831 T---GACATCATGATGACAAACAAAGTGAACACATCCATGCTCATGATGGTA 886
 Db 93087 TCAAGAAATTAAGAGGATCAAAAGAAATGGAAGACTATCCATGCTCATGGGTAGGA 93028
 QY 887 GAATCATATTTGTGAAATATGACCTATTTGCCAAAGCAATTTCAAGTTCAATGCAATTC 946
 Db 93027 GAATCAATATCTGTAATAATGCGCTACTGCGCCAGGTAATTTATGATTCAGTGCATCC 92968
 QY 947 CCACCAAAATATCATCATCTTCTTCAAGACATGAGAAAAAAACAATTTCTAAATTTCTA 1006
 Db 92967 CCATCAAGCTACCAATGATCTTTCTTCAAGATTTGAAAAAACTACTT-TAAAGTTCA 92909
 QY 1007 TGAACAACAACCAAAAAAAAAAAAAAACCAGCATAGCCAAAGCAAGCTTAGCAAAA 1066
 Db 92908 TGAACCC-----AAAAAGAGCTGATCACCAGATCATCTTAAGCATTA 92863
 QY 1067 AGAACAATCTGAGGCGATCATATCCATCTTCAAACTATATCTTCAAGGCTTAATCA 1126
 Db 92862 AGAACAAAGCTGGAGGATCATCGCTACTGCTTCAAACTATATCTTCAAGGCTTACAGTA 92803
 QY 1127 CCAAAACATCATGCGACATGACATTAACCTAGGACATAGACCAATGAGAAAAAGAGAGA 1186
 Db 92802 CCAAAACGATGTACTGTGATCCAAACAGATATAGATCAATGAAACGAACAGAGC 92743
 QY 1187 ATCCAGAAATTAAGCCAAATATATATAGCCAACTGATTTTGAACAAAGCAAAACA 1246
 Db 92742 CCTCAGAAATATATCTGATATCTAACAATATCTGATCTTTGAACAACTGAGAAAAACA 92683
 QY 1247 TAAAGTGGGAAAGACATTTCTATTAACAATGCTGATGATTTATGCGACGCATAT 1306
 Db 92682 AGCAATGGGAAAGATTCCTTATTAATTAATGATGCTGAGAAACCTGGCTACCATAT 92623
 QY 1307 GTGGAAGATGAACCTGATCCCTGCTGCTCACTTAATCAAAATTTGATACAAAGATG 1366
 Db 92622 GTAGAAAGCTGAATGATGATCCCTT-CCTTACACTTATACAAAAATTAATTAAGATGG 92564
 QY 1367 ATCAAGATCTTAATCTAGACCTTAACCAATTAATTAATTTAGAAATACATCAGAAA 1426
 Db 92563 ATTAAAGCTTAATGTTAGACCTTAACCAATTAATTAATTTAGTGAATTCCTAGGATTT 92504
 QY 1427 ATGCTTTAGACATTCCTTAGGCAAAAGATTTCAATGCGCAAGAACCAAAAGTAATGCA 1486
 Db 92503 ACCATTCAGGACATAGGCAATAGGCAAGATCTTCAATTAACCAAAAGCAATGCGCA 92444
 QY 1487 AAAAAACAAAAATTAATAGTAGGCTTAATTAATTAATTAATTTTGGCGCAAAA 1546
 Db 92443 ACAAAACCAAAATTTGACAAATGCAATCTTAATTAATTAATGACCTTCTGACAGCAAAA 92384
 QY 1547 ACATCTTAGCAGAGCAACAGCAACCAACGAGTAGAGAAATCTTCACAAATTA 1606
 Db 92383 GAAATCTACATCAGAGGAAACAGCAACCTTAACAAATGAGGAAATTTTTCGCAACCTAC 92324
 QY 1607 GCATCTGACTAAGACTTAATATCCGAATCCACAGAACTCAAAACAAATCAGCAAGAG 1666

QY 947 CCACCAATATATCATCATCTTCTTCCAGAACTAGAAAAACAATCTTAATTCATA 1006
 DB 17911 CCATCAAGCTACCAATGCTTCTTCCAGAAA-TTGGAAAACTACTTTAAAGTCCACA 17853
 QY 1007 TGGAAACAACAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1066
 DB 17852 TGGAAAC-----CAAAAAAGAGCCGATTTGCCAAGTCAATCTTAAGCCAAA 17807
 QY 1067 AGAACAATCTGGAGGATCATCATTTCCATCTTCAAACTATCTACAGGCTATATCA 1126
 DB 17806 AGAACAAGCTGGAGGATCATCATCTTCAAACTATCTACAGGCTATATCA 17747
 QY 1127 CCAAAACATCATGAGCATGATTAATAGCAATAGCAATAGCAATAGCAATAGCAAT 1186
 DB 17746 CCAAAACATGATGATCTGATGATCAAAACAGATATAGATCAATAGCAATAGCAAT 17687
 QY 1187 ATCCAGAAATTAAGCCCAATTAATTAAGCAATGATTTTGAACAAACAAAAACA 1246
 DB 17686 CCTCAGAAATTAAGCCGATATCTACAACTATCTGATCTTTGACAAACCTGAGAAAAACA 17627
 QY 1247 TAAAGTGGGAAAAAGACATCTAGTAACAAAATGCTGAGATTATTTGGCAAGCCAT 1306
 DB 17626 AGCAATGGGAAAAAGATTTCCCTATTTAATTAATGATGCTGGAAAACTGCTAGCCATAT 17567
 QY 1307 GTGGAAGATGAATCTGAATCCCTTCTCTCTCACTTAATCAAAATTTGATCAAGATGG 1366
 DB 17566 GTGGAAGATGAATCTGAATCCCTTCTCTCTCACTTAATCAAAATTTGATCAAGATGG 17508
 QY 1367 ATCAAAAGCTTAATCTGAGACTTAACATTAATAATTTAGAGATTAATCATGAGAAA 1426
 DB 17507 ATTAAGATCTTAATCTGAGACTTAACATTAATAATTTAGAGATTAATCATGAGAAA 17448
 QY 1427 ATGCTTGAACATTCATTAAGGCAAAAGATTCATGCGCAAGAACCCAAAAGTAATGCA 1486
 DB 17447 ACCATTCAGGACATAGGCAATGCGGAGGACTTCAGCTTAATAACCAAAAGCAATGGCA 17288
 QY 1487 ACAAAAACAATAATTAATAGTATAGTATTAATTAATAAATTTTGGCGAGCAAAA 1546
 DB 17387 ATAAAAACAATAATTAATAGTATAGTATTAATTAATAAATTTTGGCGAGCAAAA 17328
 QY 1547 ACAATCTTAGAGAGCAAAACAGAACCAACCGAGTGAAGAAATTTTCAAACTTA 1606
 DB 17327 GAAACTACATCAGAGAGAGAGAGCAATCAAAATGGAGAAATTTTCAAACTTAC 17268
 QY 1607 GCATCTGACTTAAGGACTTAATTCGGAATTCAGAGAACTCAAAACAATCAGCAAGAG 1666
 DB 17267 TCATCTGACAAAGGGCTAATATTCAGAAATCTAATGATGCTAATAAACAATTTACAGAA- 17209
 QY 1667 AAAGCAAAACAATCCATGAAAAGAGTGGCTTAAGGACATGAATGAATTTCAAAAGAA 1726
 DB 17208 AAAGCAAAACAATCCATGAAAAGAGTGGCTTAAGGACATGAATGAATTTCAAAAGAA 17149
 QY 1727 GATTAACAATGGCCAAACA--ACAGGAAAAAATGCTTAACATCACTAATGATTAAGGAA 1784
 DB 17148 GACTTTTATGAGCCCAAAAACATGAAAAATGCTCATATCATCTGGCCATCAAGAAA 17089
 QY 1785 ATGTAATCAACACTGTAATGCGATACCACTTATCTCTGCAAGAAATGCTCATATTTAA 1844
 DB 17088 ATGTAATCAACAACTGTAATGCGATACCACTTATCTCTGCAAGAAATGCTCATATTTAA 17029
 QY 1845 AAATCTAAAAATTAATGATGTTGGTCTGCTGCTGATTAAGAAACAATTTTCACTGC 1904
 DB 17028 AAGTCAGAGACAAACAGGTGCTGAGAAAGATGAGAAACAGAAACAATTTTCACTGT 16969
 QY 1905 TGGTGGGAATGTAATCTGGCAACAACATATGAAAAACAGTGTGAAATTTCTTAAGGAA 1964
 DB 16968 TGGGGAAGCTTAATCTGATTTCAACATTTGGAAGTCAAGTGTGCGATTTCTCAGGAT 16909
 QY 1965 CTAAAGTGAATGACCAATTTGATTCAGCAATTCCTATTA 2003
 DB 16908 CTAGAACTGGAATTAATGATTTGACCAAGCCATTCCTATTA 16870

RESULT 9
 US-10-027-983-11
 ; Sequence 11, Application US/10027983
 ; Patent No. 6617162
 ; GENERAL INFORMATION:
 ; APPLICANT: Kenneth W. Doble
 ; APPLICANT: Mark P. Roach
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
 ; FILE REFERENCE: RTS-0340
 ; CURRENT FILING DATE: 2001-12-18
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 11
 ; LENGTH: 392000
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 137740
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: unsure
 ; LOCATION: 137742
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: misc_feature
 ; LOCATION: (138122)...(138221)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: unsure
 ; LOCATION: 145507
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: unsure
 ; LOCATION: 151967
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: misc_feature
 ; LOCATION: (151967)...(154206)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: unsure
 ; LOCATION: 154217
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: misc_feature
 ; LOCATION: (164037)...(164136)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (174657)...(174756)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (186224)...(186323)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (195242)...(195341)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: unsure
 ; LOCATION: 202703
 ; OTHER INFORMATION: unknown
 ; NAME/KEY: misc_feature
 ; LOCATION: (202771)...(202870)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (206246)...(215602)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (218126)...(218225)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (220360)...(220459)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (222717)...(222816)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (223981)...(224080)
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (227487)...(227586)

OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (230157)...(230256)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (232299)...(232398)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (236552)...(236651)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (238789)...(248788)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: exon
 LOCATION: (118288)...(119101)
 OTHER INFORMATION: exon 1C
 NAME/KEY: exon:intron junction
 LOCATION: (151129)...(151130)
 OTHER INFORMATION: exon 5:intron 5
 NAME/KEY: exon:intron junction
 LOCATION: (29248)...(29249)
 OTHER INFORMATION: exon 9:intron 9
 NAME/KEY: exon:intron junction
 LOCATION: (348578)...(348579)
 OTHER INFORMATION: exon 10:intron 10
 NAME/KEY: intron
 LOCATION: (348579)...(381838)
 OTHER INFORMATION: intron 10
 NAME/KEY: intron:exon junction
 LOCATION: (386185)...(386186)
 OTHER INFORMATION: intron 11:exon 12
 US-10-027-983-11

Query Match 12.4%; Score 1068.6; DB 4; Length 392000;
 Best Local Similarity 73.8%; Pred. No. 4.7e-261;
 Matches 1515; Conservative 0; Mismatches 504; Indels 34; Gaps 11;

1 TGTATGAGCCATGCTTATATACCAAAACGAGAAAGATTA--CAAAAAGAA 58
 Db 319240 TTTATGAGCCAGCATCTCGATACCAAGCTGCGAGAGACACAAACAAAAGAGA 319239
 QY 59 ACTATAGACAGATACCATGATGATATATACATGACAGAAATCCCAACAAATATAGCTA 118
 Db 319300 ATTTTATGACCAATATCTCTGATATACATGATGACAAATCTCGATGATTAATCTGCA 319359
 QY 119 ACCCAATCCACAGCATATCAAGATATCCACCATTTGTCAAGTGGCTTCATACCA 178
 Db 319360 ACCCAATCCACAGCATATCAAGATATCCACCATTTGTCAAGTGGCTTCATACCA 319419
 QY 179 GGGTGC-AGGATAGCTTACATACACATGATCAATTAATGATACATCACTAAACGAA 237
 Db 319420 GATATCAAGGCTGCTTCAACATATGCAAGTCAATTAATCCAGCATATTAACGAA 319479
 QY 228 TTAATAACCAAAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT 297
 Db 319480 CCAATGACAAAACCAATGATATCTCAATAGATGCAAAAAAGGCTTTGACAAATCT 319539
 QY 298 AACATTTCTTATGATTAATAACCTTCAGCAAAATTCAGATAGAAGACATACCTTAATG 357
 Db 319540 AACCACTTCATGCTTAATAAAGCTCAATTAATTAAGTATGATGGAAGCTATCTCAAAA 319599
 QY 358 TTAATAAAGCCATATATGACGAGCCACAGCAAAACATTAATCTGATGAGGAAAAAGTTGA 417
 Db 319600 TTAATAAAGCCATATATGACGAGCCACAGCAAAACATTAATCTGATGAGGAAAAAGTTGA 319659
 QY 418 AAACATTTGCTTGAAGACTGGAACAGACAGCAAGATG-CTACTTTCACCACTCTATTTCA 476
 Db 319660 GAGCATTTCCCTTTGAAAACTGGGACAGAGAGGATGCTCTCTCACTCACTCTATTTCA 319719
 QY 477 ACATAGTATGAGATTTTATGACGAGCAATCAGACAGAGAAAGAAATCAAGGAGCAACC 536
 Db 319720 ACATAGTATGAGATTTTATGAGGCAATCAGGAGAGAGAAAGAAATCAAGGATTTTC 319779

QY 537 AAATCAATAAGAGAGATCAAACTGCTCCCTGTCATGATATGATATGATCTAG 596
 Db 319780 AATTAGAAAAGAGAGATCAAACTGCTCCCTGTCATGATATGATATGATCTAG 319839
 QY 597 AAAACCTTAAGACTGATCAGAAAGCTCTAGAACTGATATCAATTAATCAAGTAAAGTTT 656
 Db 319840 AAAACCTTAAGACTGATCAGAAAGCTCTAGAACTGATATGATATGATCTAG 319899
 QY 657 CAGGATCAAACTTAATGATCAAAATGATGAGTCTGATATGACCAAGTACAG 716
 Db 319900 CAGGATCAAACTTAATGATGCAAAATGATGAGTCTGATATGACCAAGTACAG 319959
 QY 717 CTGAGATCAAACTTAAGACTCAAACTTTTACATAGCT-----GTAAAAAATCACTT 771
 Db 319960 CAGAGGCCAAATCATGATGAACTTCCATTCATCAATGCTTCAAAAGAAATTAATCT 320019
 QY 772 AAGATATTTCTTACCAAGAGGTGAGAGAGCTCTACAGAGAAATCAAAACAGCT 831
 Db 32020 AAGATCAAACTTACAGAGAGGTGAGAGAGCTCTTCAAGAGAACTCAAAACAGCT 320079
 QY 832 ---GACATCATGATGACACAAACAGTGGAAACATCCATGCTCATGATGGATG 887
 Db 32080 CACAAATAAAGAGATACAAACAAATGCAAGAACATTCATGCTTATGGTGAAG 320139
 QY 888 AATCAATTTGTAAGAAATGACATATGTCCAAAAGCAATCTCAAGTCAATGCAATTC 947
 Db 320140 AATCAATTTGTAAGAAATGACATATGTCCTCAAGAGAAATCAAAATCTCAATTC 320199
 QY 948 CACCAAAATATCATCATCTTCTTCAAGAGTACAGAAATCAAAATCTCAATTC 1007
 Db 320200 CATCAAGTACCAATGATCTTCTTCAAGAA--TTGGAATAAACTCAATTAAGTTCATAT 320258
 QY 1008 GGAACAACACCAAAAAAAAAAAAAAAAAAACCAGCATGACCAAGCAAGTACGAAAA 1067
 Db 320259 GGAAC-----CAAAAAAGCCGCAATGCGCAATCATTAAGCAAA 320304
 QY 1068 GAACAATCTGAGAGGATCAGATTAACCATCTTCAACATATCTCAAGGCTATATGAC 1127
 Db 320305 GACCAAGCTGAGAGGATCAGATTAACCATCTTCAACATATCTCAAGGCTATATGAC 320364
 QY 1128 CAACAATGAGGATCTGACATTAACCTAGGACATAGCAATGAGAAAGAGAGAA 1187
 Db 320365 GAACAAGGATGATCTGATCAAAAGAGATTAAGCAATGAGAAAGAGAGAA 320424
 QY 1188 TCCAGAAATTAAGCCAAATTAATTAAGCAATGATTTTGAACAAACAAACAAAT 1247
 Db 320425 CTCAGAAATTAAGCCAAATTAATTAAGCAATGATTTTGAACAAACAAACAAAT 320484
 QY 1248 AAAGTGGGAAAAAGACATCTAGTTTAAACAAATGAGTGAATTAAGGCAAGCAGATG 1307
 Db 320485 GAATGGGAAAAAGATTCCTTATTAATTAATGATGAGGAAAACTGCTAGCAATATG 320544
 QY 1308 TGGAGAAATGAACTGATTCCTTGTCTCTCACTTAATCAAAATTTGATCAAGATGGA 1367
 Db 320545 TGAAGAGCTGAAGTCTGATTCCTTC-CTTACACTTAATCAAAATTTAATCAAGATGGA 320603
 QY 1368 TCAAGACTTAATCTGAGACTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 1427
 Db 320604 TTAAGACTTAATCTGAGACTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 320663
 QY 1428 TGCTTCAAGATCACTTGAAGCAAAAGCTTATGAGGCAAGAACCAAGTAATGCA 1487
 Db 320664 CCAATGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 320723
 QY 1488 CAAAAACAAATAATTAATGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAAT 1545
 Db 320724 CAAAAACCAATTAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 320783
 QY 1546 AACATCATTTAGAGAGCAAAAGACACACCAAGGATGAGGAAATTTTCAAACTTA 1605
 Db 320784 AGAACTTACCATCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 320843
 QY 1606 AGCATGTAGCTAAGGCTAATATCCGGAATCCAGCAAGAACTCAAAACAAATCAGCAAGAA 1665

Db	320844	CTCATTGGCAAAAGGCTATATTCAGAGATCTACAGAACTCAACCAATTTCACAGAA	320903
Qy	1666	GAAAGCAAAACAATCCCATTAAGAAGTGGGCTAAGACATGAATAGACAATTCCTCAAGA	1725
.Db	320904	AAAAACAAACACACCATCAACAAATGGGCGAAGATATGACAGACACTTCCTCAAGA	320963
Qy	1726	AGATATCAAAATGGCCAC--AAACGAGAAAAATGCTTAACATCACTAATGATTGGGA	1783
Db	320964	AGACATTTATGAGCCCAACAGACACATGAAAAATCTCATCTCCTGCGCATCGAGA	321023
Qy	1784	AATGTAAATCAACACTGTATGATGATACCACTTACTCCGCAAGATGGTGCATATTTA	1843
Db	321024	AATGCAATCAAAACCAACATGATGATACCATCTCACACCACTTAAGATGGTGAATTTAA	321083
Qy	1844	AAATCTMAAAATTAATAGATGTTGGTGGTCTGTGTGATTAAGAACAATTTCACATG	1903
Db	321084	GAAATCAAGGAAAAACAGGTGCTGGAGAGAGATGTGAGAAATGGAAGACTTTTACATG	321143
Qy	1904	CTGGTGGGAATGTAAACTTGGCCGAACCACTATGAAAAACAGTGTGAAAAATTTCTTAAGA	1963
Db	321144	TGGTGGGACTGTAAACTAGTTCAACCAATGTGGAATTCAGTGTGGCAATTCCTCAGGGA	321203
Qy	1964	ACTAAAGATGATCCAGCACTTTGATCCAGCAATCCCATTAATATGTATTAATATATATA	2023
Db	321204	TCTAAGAACTGAAATACCACTTTGACCCAGCCATCCCAATT-ACGGGTATATACCCAAAG	321262
Qy	2024	TTTATATACCATG 2036	
Db	321263	ATTATTAATCATG 321275	

```

RESULT 10
US-10-071-411A-63/c
; Sequence 63, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Glenn Barnes
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; TITLE OF INVENTION: 5-Lipoxygenase Gene
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/071,411A
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 168174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)_(168174)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411A-63

```

Query Match	12.3%	Score 1064.6	DB 4	Length 168174
Best Local Similarity	74.0%	Pred. No. 2.9e-260		
Matches 1519	Conservative	0	Mismatches 499	Indels 35
				Gaps 12

Qy 1 TGTATGAGGCCATGTCATTATATCCGAAACAG--AAAGGATATCAAAAAAAAAA 58
Db 122448 TTTATGAGGCCGATCATCTCTGATACAAAGCTTGCGAGTACACACAAAAAAAAACA 122489
Qy 59 ACTATAGACCAATATACATGATGATATATACAGCAGAAATCCCACAAAATCTAGCTA 118
Db 122468 ATTTTAGACCAATATCTCTGATGATCATGATGCAAAAAATCTCTCATPAAATATCTGCA 122429
Qy 119 ACCCAATCAACAGCATATCAAGAAATATCAATCCATCTGTCAAGTGGGTTTCATACAG 178

Db	122428	ACCGAATCCAGACGACATCAAAAAAGCTTATCCACCGAGATCAAGTGGGCAATATCCCTG	122365
Qy	179	GGGTGC-AGGATAGGTTAACATACAAAGTCAATAATGTGATACATCAATTAACAGAA	237
Db	122368	GGATGCCAAAGCTGTTTCAACATACAAATCAATAATGTAAATCCAGCATATTAACAGAA	122308
Qy	238	TTAAAAACAAAAATCAGATGATCATCTCAATGATGCTGAAAAAGATTTTGACAAATCT	297
Db	122308	CCAAAGACAAAAACACATGATTAATCTCAATGATGACAAAAAGGCTTTGACAAATTC	122249
Qy	298	AACATTTCTTATGATTTAAAACTTACGCAAAATGACATAGAAAGGACATACCTTAATG	357
Db	122248	AACAAACATTTATGCTTAAAAATCTCTCATTAATTTAGATTTGATGGAGCTATCTCAAAA	122188
Qy	358	TAAATAAAGCCATATATATGACGAGCCACACAGCAAACTTTACTGATTTGGGAAAAGTTGA	417
Db	122188	TAAATAGAGCTATTTATATGCAAAACCAACAAATATCATATCTGAATGGGCAAAAATCTGG	122129
Qy	418	AAACATTTGTCCTTGGAACCTGGAAACAGACAGAGATG-CTACTTTTACCACTTCTATTCA	476
Db	122128	AAGCATTTCCCTTTGAAAACTGGCAACAGACAGAGATGCCCCCTCTCTCACCACTCTTAATCA	122068
Qy	477	ACATAGTAGTGGAGGTTTATAGCCAGAGCAATGACAAAGAGAAAGAAATCAAGGGACCC	536
Db	122068	ACATAGTGTGGAGGTTCTGGCCAGGGCAATGACGAGAGAAAGAAATTAAGGGATTT	122008
Qy	537	AAATCAATTAAGAGAGATCAAACTGTCCTCTGTTCACTGATGATATGATTTGATTAACCTAG	596
Db	122008	AATTAGAGAAAGAGAGATCAAAATTTGTCCTCTGTTCCAGATGACATGATTTGATTAATTTAG	121949
Qy	597	AAACCCCTTAAGACTCATCTCAGAAAGCTCC-TAGAACTGATACATTAATTCAGTAAAGTT	655
Db	121948	AAAACCCCATTTGCTCAGCCCAAAATCTCTTAAAGCTGATAGGCAATTCAGCAAAAGTC	121889
Qy	656	TCAGAGTACAACTTAATTTGACAAATACAGTACAGCTGTTACACCAACGTGACCAA	715
Db	121888	TCAGAGTACAAAAATCAATGTCAGAGATCAAGAAATCAAGCAATCTTATACCAAAATACAGACAA	121829
Qy	716	GCTGAGAAATCAAAATCAAGAACTCAAAACCTTTTAACTAGCTGTAAA-----AAATATAC	769
Db	121828	ACAGAGAGCCAAATCAGAGATGAATCTCCATTCACAAATTTGCTACAAAGAGAAATTAATAC	121765
Qy	770	TTAAGAAATATTTCTTACCCAGAGAGGTGAAGACCTCTACAGAGAAAATCTACAAAACACAG	829
Db	121768	CTAAGAAATTCAACTTACAAAGGATGTAAGGACCTCTTCAAGAGAAATCTACAAATCAGCTG	121709
Qy	830	CT-----GACATCATATGATGACACAAACAGTGGAAAACATCCCATGCTCATGATGGGT	885
Db	121708	CTCAACGAATTAAGAGAGACCAAAACAAATGGAAGAACATCCATGCTCATGAGATAGGA	121649
Qy	886	AGAACTCAATATTTGTGAAAATGACATATTTGCCAAAAGCAATCTACAAAGTTCAATGCAATT	945
Db	121648	AGAAATGAATATATTAATAAATGGCCATACT-CCCAAGATTAATTTATAGATTCAATGTCAATC	121590
Qy	946	CCCAACAAAATATCATCATCATCTTCTTACAGAACTAGAAAAAAACAAATTTCTAAATTTCAAT	1005
Db	121589	CCCATCAAACTACGAATAGACTTTCTTCCAGAA-TTGGAAAAAATCTAGCTTAAAGTTCAAT	121531
Qy	1006	ATGGAACACAAACAAAAAATTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	1065
Db	121530	ATGGAAAC-----CAAAAAAGAGCCGCAATGGCCMAAGCAATCTCTAACCCAA	121485
Qy	1066	AAGAACAAATCTGAGAGCATCATTTACCCATCTTCAAACTATATCTACAAAGGCTATATATC	1125
Db	121484	AAGAACAAAGCTGAGAGCATCATGCTACTGACTTCAATCTATATCTACAAAGGCTACAGTA	121425
Qy	1126	ACCAAAACATCATGGCACTGACATTAATACTAGGCAATATAGACCAATGAAAAAGAGAG	1185
Db	121424	ACCAAAACAGCATGTGTAAGTGAACAAAACAGAGATATAGACCAATGGAACGAACAGAG	121365
Qy	1186	AATCGAAGATTAAGCCAAATATTAATTAAGCCAACTGATTTTGGACAAAGCAAAACAAAC	1245

Db 121364 CCTCAGAAATGACACACACATCTACACATCTGATCTTTGACAAACCTGACAAAC 121305
 Qy 1246 ATAAAGTGGGAAAAGACATTTAGTAAACAAATGGTCTGAGATTATTTGGCAAGCCACA 1305
 Db 121304 AAGCATGGGAAAGATTCCTATTTAAATGATGGGAAAACCTGGCTAGCCATA 121245
 Qy 1306 TGTGAGAGATGAACTGATCCCTTGTCTCTCACTTAATACAAAATTTGATCAAGATG 1365
 Db 121244 GGTAGACAGCTGAGCTGATCCCTTC-CTTACACCTTTATACAAAATTTATTCAGAGTG 121186
 Qy 1366 GATCAAGACTTAATCTGAGACTTAAACCATTAATTTCTAGAGATTAATCAAGAA 1425
 Db 121185 GATTAAAGCTTAATTTGATGACTTAAACCATTAACCACTTGAAGAAAACCTTAGCAA 121126
 Qy 1426 AATGTTCTAGACATTCATTTAGGCAAGACCTTCAAGCCACAAACCCAAATTAATGC 1485
 Db 121125 TACCATTCAGACATTAAGATGGGCAAGACTTCAATGCTTAAACACCAAAACCAATGG 121066
 Qy 1486 AACAAAACAAAATTAATAGATAGAGCTTAATTAATTAAGCTTTTGGCCAGCCAAA 1545
 Db 121065 AACAAAGCCAAATAGACAAATGGGATCTTAATTAATTAAGGCTTCTGCAACGCAAA 121006
 Qy 1546 AACAACTATTAGCAGAGCAACAGACACACCCAGGAGAGAGAAAATTTTCAAACTA 1605
 Db 121005 AGAACTACATCAGAGTGAACAGGCACTTACAGATGGGAAAATTTTCAATCTA 120946
 Qy 1606 AGCATCTAGCTAAGACTAATATCCGGAATCCCAAGAACTCAAAACAAATCGCAAGAA 1665
 Db 120945 CTCACTGCAAAAGGGCTAATTCAGAAATCTAACAACCACTCAAAATTTTACAGAA 120886
 Qy 1666 GAAAGCAAAACATCCCATGAAAGAGTGGCTAAGAGACATGAATAGCAATTTCTCAAAAG 1725
 Db 120885 AAAAACAACCAACCCCATCAAAAAGTGGCGAAGATTAAGACACCTTCTCAAAAG 120826
 Qy 1726 AGATATACAAATGCCCCAAC-AAAACGAAAAAATCTTAATCACTAATGATTAAGGA 1783
 Db 120825 AGACCTTTATGACAGCCACAGACACATGAATAAATCTCATCTGCGCCATCAGAGA 120766
 Qy 1784 AATGAATCAACACCTGAATGATGATACCACTTCTCTGCAAGAAATGCTAATTTA 1843
 Db 120765 AATGCAATCAAAACATATATGATATACATCTCACACAGTTACATGCAATCTA 120706
 Qy 1844 AAAATCTAAAAATATAGATGTTGGTCTGCTGCTGATTAAGAAACCTTTTACACTG 1903
 Db 120705 AAAGTCAAGAACACAGGTGCTGAGAGGGTGTGAGAAATGGGAACTTTTAACTG 120646
 Qy 1904 CTGGTGGGAATGTAATCTTGGCAACCACTATGAAAAACGTTGGAATTTCTTAAGA 1963
 Db 120645 TTGGTGGGACTGTAACTGATTCACCTTGTGGAAGACACTGTGGCAATTTCTCAGGG 120586
 Qy 1964 ACTAAAGATGATCGACATTTGATCGAGATCCCATTAATATGATTAATATATATA 2023
 Db 120585 TCTAGAACTAGAAATACCAATTTGACCCAGCCATCCATTT-ACTGGATATATACCAAAG 120527
 Qy 2024 TTTATATACCATG 2036
 Db 120526 ATTATATATCATG 120514
 RESULT 11
 US-10-071-411A-2/c
 ; Sequence 2, Application US/10071411A
 ; Patent No. 6797475
 ; GENERAL INFORMATION:
 ; APPLICANT: Joanne Meyer
 ; TITLE OF INVENTION: Detection of Polymorphisms in the Human
 ; FILE REFERENCE: MRI-021
 ; CURRENT APPLICATION NUMBER: US/10/071,411A
 ; PRIOR FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: 60/267,515
 ; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/314,248
 ; PRIOR FILING DATE: 2001-08-21
 ; NUMBER OF SEQ ID NOS: 66
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 168273
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(168273)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-071-411A-2
 Query Match 12.3%; Score 1064.6; DB 4; Length 168273;
 Best Local Similarity 74.0%; Pred. No. 2.9e-260;
 Matches 1519; Conservative 0; Mismatches 499; Indels 35; Gaps 12.
 Qy 1 TGTATGAAAGCAATGCTACTTTAATACCAAAACAGG--AAAGATATACAAAAGAA 58
 Db 122647 TTTATGAGGCCAGCATCATCTGATACCAAGCTGGAGATACACACAAAACAA 122588
 Qy 59 ACTATGACAGTACCTGATGAAATATATCATGCAAAATCCCAACAAATCTAGCTA 118
 Db 122587 ATTTTGAACCAATATCTCTGATGAACATGATGCAAAATCTCAATAAATCTGGCA 122528
 Qy 119 ACCCATCAACAGCATATACAAAGATATCCATTTGCAATGGGTTTCATACAG 178
 Db 122527 ACCGAATCAGACGACATCAAAAGCTTATCCACAGATCAATGGGATCATCTCTG 122468
 Qy 179 GGGTGC-AGATATAGTTTAAATATACACAAATCAATTAATGATATCATATCAATTAACAGAA 237
 Db 122467 GATGGAAGCTGGTTCAATACATACCAATATATATATATCAAGATATTAACAGAA 122408
 Qy 238 TTAATAAACAATAATCATGATCATCTCATATGATGCTGAAAAGCAATTTGACAAATCT 297
 Db 122407 CCAAAAGCAAAACCAAGATTAATCTCAATGATGCAAAAAGGCTTTGCAAAATTC 122348
 Qy 298 AACATTTCTTTATGATTAATTAACCTTACAGCAAAATGACATATGAAGGACATACCTTAATG 357
 Db 122347 AACACCATTCATGCTTAATAACCTCTCATATTAATTAATGATGATGAGACATCTCAAA 122288
 Qy 358 TATATAAGCCATATATGACGACACCAAGCAAACTTAATCTGAATGGGAAAAGTTGA 417
 Db 122287 TATATAGACTATTTATGACAAACCAACCAATATCATCTGAATGGGCAAAATCTG 122228
 Qy 418 AAACATTTGCTCTGAGAAATGGAACCAAGCAAGATG-CTACTTTCACTCTTATTC 476
 Db 122227 AAGCATTCCTTTGAAATCTGCAACAGACAGGATGCCCTCTCACACATCTCTATTTCA 122168
 Qy 477 ACATAGTAGTGAAGTTTATGACGACGCAATGACACAAAGAAATCAAGGGGACCC 536
 Db 122167 ACATAGTGTGGAAGTTTCTGGCAGGGCAATGAGGAGAAAGAAATTAAGGATATT 122108
 Qy 537 AAATCAATTAAGAGAGAGTCAAACTGCTCCCTGTTCACTGATGATATGATTAATCTAG 596
 Db 122107 AATTAGAAAAGAGAGAGTCAAAATGCTCTGTTTCAAGATGATGATTAATTTAG 122048
 Qy 597 AAAACCTTAAGATCTATCAAGAAAGCTCC-TAGAACTGATATCAATTAATTCATGTAAGTT 655
 Db 122047 AAAACCCCATTTGCTCAGCCCAAAATCTCTTAAGCTGATAGGCACTTCAGCAAACTC 121988
 Qy 656 TCAGGATCAAACTAAATGATACAAATTCAGTACGCTATACCAACAGTACGCA 715
 Db 121987 TCAGATACAAATCAATGATGCAAGATTCAGAAATCTTATACCAATTAACAGACAA 121928
 Qy 716 GCTGAGATCAATCAAGAACTCAAACTTTTAAATAGCTGTAA-----AAAATAC 769
 Db 121927 ACAGAGGCAATTCAGAGTGAATCTCCATTTCAAAATTTGCTACAAAGAGATTAATATAC 121868
 Qy 770 TTAAGAAATTTCTTACCAAGAGGATGAAGAGCTCTACAGAGAAATCAACAAACAG 829
 Db 121867 CTAGAAATCAACTTACAAAGGATGTAAGAGCTCTTCAAGAGAACTACAAATCACTG 121808

QY 830 CT-----GACATCATAGTACAAAGAGTGAAGACATCCCATGCTCATGATGGT 885
 DB 121807 CTCAACCAAAATTAAGAGACAAACAAAGAGAAACATCCCATGCTCATGATGAGA 121748
 QY 886 AGAATCAATATGTAAGAAATGACATATGCGCAAGCAATCTCAAGTTCAATGCAATT 945
 DB 121747 AGAATGATATCATTAATAATGCGCAATCT -CCAAAGGTAATTTATGATTAATGTCATC 121689
 QY 946 CCCACCAAAATATCATCATCTTCTCAAGACATGAAAAAAACAAATTTCTAAATTCAT 1005
 DB 121688 CCATCAAACTACCAAGACCTTTCTCTCAGAA -TTGGAAAAATCTACGTTAAAGTTTAT 121630
 QY 1006 ATGGAACCAACACCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
 DB 121629 ATGGAAC-----CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121584
 QY 1066 AGAACAATATGAGGATCATCTTCCCATCTTCAAACTATCTCAAGGCTATATC 1125
 DB 121583 AGAACAATATGAGGATCATCTTCCCATCTTCAAACTATCTCAAGGCTATATC 121524
 QY 1126 ACCAAAAATCATGAGCATGACATTAACATGAGACATGACATGAGAGAGAGAGAGAG 1185
 DB 121523 ACCAAAAATCATGAGCATGAGTACCAAAACAGAGATATGACCAATGAGAGAGAGAGAG 121464
 QY 1186 AATCGAATTAAGCCCAATTAATTAAGCCATGATTTTGAACAAAGCAACAAAC 1245
 DB 121463 CCCTCAAAATTAAGCAACATCTACCAACCATCTGATCTTTGAACAAACGTAACAAAC 121404
 QY 1246 ATAAAGTGGGAG 1305
 DB 121403 AAGCAATGGGAG 121344
 QY 1306 TGTGGAAGATGAAG 1365
 DB 121343 GGTAG 121285
 QY 1366 GATCAAGATTAATCTGAGACCTTAACCAATTAATTAATTAATTAATTAATTAATTA 1425
 DB 121284 GATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 121225
 QY 1426 AATGCTTCTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1485
 DB 121224 TACCAATGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 121165
 QY 1486 AACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1545
 DB 121164 AACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 121105
 QY 1546 AACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1605
 DB 121104 AGAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 121045
 QY 1606 AGCAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1665
 DB 121044 CTCAATGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120985
 QY 1666 GAAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1725
 DB 120984 GAAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120925
 QY 1726 AGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1783
 DB 120924 AGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120865
 QY 1784 AATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1843
 DB 120864 AATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120805
 QY 1844 AATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1903
 DB 120804 AATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120745

QY 1904 CTGATGGAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1963
 DB 120744 TTGATGGAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120685
 QY 1964 ACTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2023
 DB 120684 TCTGAACTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120626
 QY 2024 TTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2036
 DB 120625 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120613
 RESULT 12
 US-09-816-095-3/c
 ; Sequence 3, Application US/09816095
 ; Patent No. 6664084
 ; GENERAL INFORMATION:
 ; APPLICANT: GAN, Weiniu
 ; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES.
 ; FILE REFERENCE: C1001147
 ; CURRENT FILING DATE: 2001-03-26
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 9916
 ; TYPE: DNA
 ; ORGANISM: Human
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(9916)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-816-095-3
 Query Match 12.3%; Score 1064.2; DB 4; Length 9916;
 Best Local Similarity 74.0%; Pred. No. 2.6e-260;
 Matches 1493; Conservative 0; Mismatches 488; Indels 36; Gaps 10;
 QY 1 TGTATGAGCAATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 58
 DB 21190 TTTATGAGGCGAGCATATCTCTGATTAACCAAGCGGCGAGAGACAGCCAAAAAGAGA 21131
 QY 59 ACTATGACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 118
 DB 21130 ATTATGACCAATATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 21071
 QY 119 ACCCAATCAACAGCATATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 178
 DB 21070 ACCCAATCAACAGCATATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 21011
 QY 179 GGGTGC-AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 237
 DB 21010 GGAATGCAAGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 20951
 QY 238 TTAATAAACAATAATCAATGATCATCTCAATAGATGCTGAAAAAGATTTGAACAAATCT 297
 DB 20950 CCAAGCAAAAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 20891
 QY 298 AACATTTCTTATGATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 357
 DB 20890 AACCAAGCTTATCTTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 20831
 QY 358 TAAATAAAGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 417
 DB 20830 TAAATAAAGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 20771
 QY 418 AACATTTCTTATGATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 476
 DB 20770 AGGATTTCTTATGATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 20711

QY	4777	ACATGTAGTGGAAAGTTTTCGCGACGAATCAGACAAAGAAAGAAATCCAGGGCACCC	536
Db	20710	ACATGTGTGGAAAGTCTTGCCCGAGGGGCATTTAGCGAAGAAAGGACATTAAGGGTATTC	20651
QY	537	AAATCAATTAAGAGGAAGTCAAACTGTCCCTGTTCACTGATGATVTGATGTATCTAG	596
Db	20650	AATTAGGAAAAGAGGGAGTCAAAATGTGCCGTGTTTGACAGACGACATGATTTGATATCTAC	20591
QY	597	AAAACCTTAAGAAGTCACTCCAAAGAGCTCTGAACTGATACATTAATTCAGTAAAGTT	656
Db	20590	AAAACCCCAATGTCTCAGCCAAAATCTCCTTAAGCTGATTAAGCAATTCAGCAAGCTCT	20531
QY	657	CAGGATACAAACTTAATGTACAAATTCAGTACGCTCTTATACCCAAAGTACCAAG	716
Db	20530	CAGGATACAAAATCAATGTACAAAATTCACAGGACATTTATATACCAATTAACAGACAA	20471
QY	717	CTGAGATCAAAATCAAGAACTCAACACTTTTACATAGCT-----GTAAAAAATCT	770
Db	20470	CAGAGAGCCAAATCATAGAGTAATCCCATTCACAAATTGCTTCAAGATTAATTAATCT	20411
QY	771	TAAGAATATTCTTACCCAGAGGTGAAGACTCTTCAAGAGAAACTACAAACACAGC	830
Db	20410	TAGGAATCCAACTTACAAAGGAGCTGAAAGGACCTTTCAAGAGAACTACAAACCACTGC	20351
QY	831	-----TGACATCATAGATGACACAAACAAGTGGAAAACAATCCAGCTCATGGATGGGTA	886
Db	20350	TCAATTAATTAAGAAAGATACAAACAATTAAGAAACATTCCTCATGGCTCATGGTAGGA	20291
QY	887	GAATCAATATTGTGAAGAAATGACATATTGCGAAAAAGCAATCTACAGATTCAATGCATTC	946
Db	20290	GAATCAATATTGTGAAGAAATGGCCATATCTGCCAAGTAAATTTATGATTTCAATGCCATCC	20231
QY	947	CCACCAAAATATCATCATCTTCTTTCACAGACTAGAAAACAAATCTTAATTTCTATA	1006
Db	20230	CCATCAAGCTACCAATGACTTTCTTCACAGAA--TTGAAAAAACTACTTTAAAGTTTATA	20172
QY	1007	TGGAAACAACACCAAAAAAABAAAAAACCCTGATGCCAAAGCAGACTTACCAAA	1066
Db	20171	TGGAAAC-----CATTAAGAGGCCACATGTCAGAGTCAATCTCTAACCCAAA	20126
QY	1067	AGAACCAATCTGAGGCGATCACAATTCCACTTTCMACTATATCTATACAGGCTTATATCA	1126
Db	20125	AGAACCAAGCTGAGGCGATCACACTACTGACTTCAAACTATATCTACAGGCTTACGTA	20066
QY	1127	CCAAAAATCATGSGCACTGACATTAACCTAGGCACTATAGCCCAATGAAAGAAAGAGAGA	1186
Db	20065	CCAAAAACAGCTGTACTGTGACCAAAACAGAGATATGATCAATGGAACAGAACGAGGC	20006
QY	1187	ATCCGAAATTAAGGCCAAATTAATTAATATAGCCACTGATTTTTCACAAAGCAACAAAAACA	1246
Db	20005	CCTGCAAAATTAAGCGCTGCATATCTACAACTATCTGATCTTTTTCACAAACCTGAAAAAACA	19946
QY	1247	TAAAGTGGGGAAGAAGACATCTAGTTTAACAATAGGTGCTGAGATTAATTGGCAAGCCACAT	1306
Db	19945	AGCAATGGGGAAGAAGATTCCCTATTTAATTAATAGTGTGGGAAAACTGGCTAGCCATAT	19886
QY	1307	GTGGAAGAAATGAAGCTGGATCCCTGTCTCTCACTTAATACAAAAATTGATCAAGATGG	1366
Db	19885	GTAGAAAGCTGAAGCTGATCCCTT--CCTTACACTTATACAAAAATTAATCAAGATGG	19827
QY	1367	ATCAAAAGCTTAATCTGAGACTTAACCAATAAAATTTCTAGAAATTAATCATCAGAAA	1426
Db	19826	-----AAACTTAAACGTTAGACTTAAACCAATTAAMCCCTAGAAAGAAACCTTAGGCAT	19772
QY	1427	ATGCTTCTAGACATCTACTTAGGCAAAAGACTTCAATGGCCAGAAACCCAAAGTAATGCA	1486
Db	19771	ACCATTCAGACATAGGCAATGGGCAAGGACTTCATGCTTAATAACCCAAACCATTTGGCA	19712
QY	1487	ACAAAAACAATAATTAATAGTAGACTTAATTAACCTTAATAAGCTTTTTCGACACAAA	1546
Db	19711	ACAAAAGCCAAATTTGACAAATGGGATCTTAATTAACCTTAAGAGCTTTCTGCACACAAAA	19652
QY	1547	ACATATTTAGCAGAGCAACAGACAAACCCAGGTGAGAGAAATCTTCACAAACTTA	1606

Db	1951	GAAACCTACATCAGAGTGAACAGGCACCTTACAAATGGAGAAATATTTGCACTTAC	19592
Qy	1607	GCATCTGCACTAAGACTAAATATCCGGAATCCACAAGAACTTCAAACAAATCAGCAGAAG	1666
Db	19591	TCATCTGACAAAGGGCTAATATTCACGAATCTTACATGAATCTCAACAAATTTTACAGAAA	19532
Qy	1667	AAAGCAAAACATTCCTCGTGAAGAAGTGGGCTAAGGACATGAATATGACAAATTCCTCAAAAGAA	1726
Db	19531	AAACCAACCAACCCCATCAACAAGTGGGGAAGACATGAATATGACACTTCTCAAAAGAA	19472
Qy	1727	GATATCAAAATGGCCCAACAAGGAAAAATATGCTTAACATCACTTAATGATTAAGGGAAT	1786
Db	19471	GACATTTATGCAAGCCCAAAAGACATATGAAAAAAC-GCTCATCATCTGGCCATCAGAGAAAT	19413
Qy	1787	GTAATCAACACTGTATATGCGATACCACTTACTCTTGCAAGAAATGGTCATATTTAAAA	1846
Db	19412	GCAATTCAAAACCAAGATGAGATACCATCTCAACACATTTAGAAATGGTATCTAAAAA	19353
Qy	1847	ATCTAAAAATATATAGATGTTGGTGGTCTGTGGTATATAAGAACACTTTTACCTGCTG	1906
Db	19352	GTCAGGAAACAAACAGGTGTCTGAGAGAGATGTGAGAAAAATAGAACACTTTTACCTGTG	19293
Qy	1907	GTCGGAATGTAACTTGCCCAACAACATAATGAAAAACATGTGTGAAAAATTTCTTAAGAACT	1966
Db	19292	GTGGGACTGTAACTGTAAATGTAACCAATGTGTGAAATCAATGTGGGCAATTCCTCAGGAACT	19233
Qy	1967	AAAAGTAGATCGACATTTGATTCGACCAATCCGATTA	2003
Db	19232	AGAACTAGAAATATCCATTTGATTCAGCCATCCCATTA	19196

RESULT 13

US-10-071-411A-63

; Sequence 63, Application US/10071411A

; Patent No. 6797475

; GENERAL INFORMATION:

; APPLICANT: Glenn Barnes

; APPLICANT: Joanne Meyer

; TITLE OF INVENTION: Detection of Polymorphisms in the Human

; FILE REFERENCE: MRI-021

; CURRENT APPLICATION NUMBER: US/10/071,411A

; CURRENT FILING DATE: 2002-02-07

; PRIOR APPLICATION NUMBER: 60/267,515

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/314,248

; PRIOR FILING DATE: 2001-08-21

; NUMBER OF SEQ. ID NOS: 66

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 63

; LENGTH: 168174

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(168174)

; OTHER INFORMATION: n = A,T,C or G

US-10-071-411A-63

Query Match	12.3%;	Score 1058.6;	DB 4;	Length 168174;
Best Local Similarity	73.4%;	Pred. No. 9.8e-259;		
Matches 1526;	Conservative 0;	Mismatches 514;	Indels 39;	Gaps 12;

1 TGTATGAAGCCATGTCACTTTAATATCCAAAAACAGAAAGATATTA--CAAAAAAGAAA 58

TTTATGAGGCCACATCATCTCTGATDCCAAAGCCGGCAGAGACAAAACCAAAAAAGG 159587

ACTATAGACCAAGTACCATGATGAATATATACATGCAGAATCCCAACAAATACTAGCTA 59

ATTTTAGACCAATATCTCTTGAAGAACATTTGATGCAAAATCTCTCAATTAATACTGGCAA 159647

ACCAATCCAAACGATATCAAGAAATATATTCACCATTTGTCAAAGTGGTTTCATTCACAG 119

Db 159707 ACCAAATCCAGCAGCACTCAAAAAGCTATCCAGCATGATCAAGTGGCTTCACTCCCTA 159766
 Qy 179 GGGTGC-AGGATAGGTTAACTACATACAGCAAGTCAATTAATGATGATCATCATATAACAGAA 237
 Db 159767 GGATGCAAGGCTTGTCAAGTATGCAAAATCAATTAATGATGATCAAGATTAATTAACGAA 159826
 Qy 238 TTAATAAACAATAATCATGATCATCTCATATAGTCTGAAAAAGCATTTGACAAATCT 297
 Db 159827 CCAAAAGCAAAAACCATATGTTATCTCAATATGATGAGAAAAGCCCTTTGACAAATTC 159886
 Qy 298 AACATTTCTTATGATTAATTAACCTTCAGCAAAATGACATGAGAAAGACATACCTTAATG 357
 Db 159887 AACAACTTCATATGTTAAAACTCTCAATTAATGATGATGCTGGAGCTATCTCAAAA 159946
 Qy 358 TAAATAAAGCATATATGACGAGCCACAGCAAACTTATATCTAGATGGGAAAAAGTTGA 417
 Db 159947 TAAATAGAGCTATATGTCAAACCCACAGCCATATCATGATGGGAAAAAGCTGG 160006
 Qy 418 AAACATTTGCTTGAAGACTGGAACAAGACAAAGATG-CTACTTTCACCACTTCTATTC 476
 Db 160007 AAGCATTTCCCTTGAATAATGGCAACAAGAGGATGCTCTCAGCACTCTCATTTCA 160066
 Qy 477 ACATAGTAGGAGGATTTTGAAGCAATCAAGCAAGAAAGAAATCAAGGCAACC 536
 Db 160067 ACATAGTGTGGAAGATCTGTCAGGCAATCAGGAGAGAAAGAAATTAAGGGATTC 160126
 Qy 537 AAATCAATTAAGAGAGAGTCAAACTGCTCCCTGTTCACTGATGATGATGTTATATCTAG 596
 Db 160127 AATTAGAAAAAGAGAGTCAAAATGCTCCCTGTTGAGATGATGATTTTATATCTAG 160186
 Qy 597 AAAACCTTAAGACTCATCCAGAAAGCTCTTGAAGCTGATATCAATTAATTCAGTAAAGTT 656
 Db 160187 AAAACCCCATTTGTCAGCCCAAAACCTCTTAACTGATTAAGCACTTCAGCAAGTCT 160246
 Qy 657 CAGGATCAAACTTAATGATGACAAATCAAGTACGCTTATACCAAGCAAGTACCAAG 716
 Db 160247 CAGGATCAAAATTAATGATGACAAATCAAGTATCTTATTAACAACAAGCAAGCAAA 160306
 Qy 717 CTGAGATCAAAATCAAGAGTCAAACTTTTACATAGTGTAA-----AAATATCT 770
 Db 160307 CAGAGACCAATATATGATGATCTCCATTCACATTTGCTTCAAGAGATTAATTAATCC 160366
 Qy 771 TAAAGATATTTCTTACCAAGAGGATGAGAACTCTTCAAGAGAAATCTCAAAACACAGC 830
 Db 160367 TAGGAATCCACCTTAACAAGGATGAGAACTCTTCAAGAGAACTCAAAACACAGC 160426
 Qy 831 T-----GACATATGATGACCAAAAGTGAAGAACTCCATGCTCATGATGGGTA 886
 Db 160427 TCAACGAATTAAGAGAGATCAAAACAAATGAGAACTTCATGCTCATGGGTAGGAA 160486
 Qy 887 GAATCAATATTTGAAAAATGACCATTTGCCAAAGCAATCTACAAAGTTCATGCAATTC 946
 Db 160487 GAATCAATATTTGAAAAATGACCATTTGCCAAAGTTCATGCAATTCATGCTCATTC 160546
 Qy 947 CCAACCAAAATATCATCATCTTCTTCAAGAACTGAAAAAAACATTTCTAAAAATTCATA 1006
 Db 160547 CCATCAAGCTACCAATTAATCTTCTTCACTAA-TTGAAAAAACTACTTTAAAGTTCA 160605
 Qy 1007 TGGACACACACCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1066
 Db 160606 TGGAGC-----CAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 160651
 Qy 1067 AGAACAATCTGAGGAGTCACTTACCATCTTCAAACTATGATGATGATGATGATGATGATG 1126
 Db 160652 AGAACAAGGCTGGAGGAGTCACTTACCATCTTCAAACTATGATGATGATGATGATGATG 160711
 Qy 1127 CCAAAACATCATGCACTGACATTAAGCACTAGCACTAGCACTAGCACTAGCACTAGCA 1186
 Db 160712 CCAAAACAGCTTGTACTGTTACCAAAACAGAGATATGATCAATGAGAAAGAGAGAGC 160771
 Qy 1187 ATCCAGAAATTAACCAATTAATTAAGCCTGATTTTGAACAAAGCAACAAACA 1246

Db 160772 CCTCAGAAATTAATGCTGATATGACAACTGTCTGATCTTTGACAAACCTGAGAAAAA 160831
 Qy 1247 TAAAGTGGGAAAAAGACATTTCTAGTTAACAAAATGCTGAGATTAATTTGGCAACCAT 1306
 Db 160832 AGAAATGGGAAAAAGATTTCCCTATTTAATTAATGCTGAGAAAACTGGCTAGCCATAT 160891
 Qy 1307 GTGGAAGATTAAGAACTGATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1366
 Db 160892 GTAGAAAGATTAAGAACTGATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 160950
 Qy 1367 ATCAAGATCTTAATCTGAGACCTTAACCAATAAAATTTCTAGAAATTAACATCAGAAAA 1426
 Db 160951 ATTAAGATCTTAATCTGAGACCTTAACCAATAAAATTTCTAGAAATTAACATCAGAAAA 161010
 Qy 1427 ATGCTTGAACATTCATTAAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1486
 Db 161011 ACATTCAGGAGCANTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 161070
 Qy 1487 AAAAAAACAATAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 161071 AAAAAAACAATAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 161130
 Qy 1547 ACATTCATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1606
 Db 161131 GAATCTACATCAGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 161190
 Qy 1607 GCATCTGATTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1666
 Db 161191 TCATCTGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 161250
 Qy 1667 AAGCAAAACAATCCCATGAGAAAGTGGGCTTAAGGACATGATGATGATGATGATGATGAT 1726
 Db 161251 A-----AAACAAACCCCATCAAAAGTGGGCAAAATGATGATGATGATGATGATGAT 161306
 Qy 1727 GATTAACAATGAGCAAC--AAACAGAAAAAATGCTTAACATGATGATGATGATGATGATG 1784
 Db 161307 GACATTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 161366
 Qy 1785 ATGTAATTAACAACATGTAATGCAATACCACTTATCTCTGCAAGAAAGTGCATTAATTA 1844
 Db 161367 ATGTAATTAACAACATGTAATGCAATACCACTTATCTCTGCAAGAAAGTGCATTAAT 161424
 Qy 1845 AAATCTAAAAATTAATGATGTTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1904
 Db 161425 AAGTCAGAAACAACAGATGCTGAGAGGACATGAGAAATGAGAAATGCTTTTACACTGT 161484
 Qy 1905 TGGTGGGAATGTAATCTTGGCAACCACTATGAAAAAGTGGGAATTTCTTAAGGAA 1964
 Db 161485 TGGTGGGAGGTAAATTAATGTTCAATCAATGTTGAAAGAGTGGCAATTTCTCAAGGAT 161544
 Qy 1965 CTAAGATGATGACCATTTGATCCAGCAATCCCATTTAATATGATTAATATATATATAT 2024
 Db 161545 CTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 161603
 Qy 2025 TTATATACATGAGATCAACTGACCACTAATAAAAAAGAT 2063
 Db 161604 TTACAAATGAGCACTATTAAGACATGACACAGAT 161642

RESULT 14
 US-10-071-411A-2
 ; Sequence 2, Application US/10071411A
 ; Patent No. 6797475
 ; GENERAL INFORMATION:
 ; APPLICANT: Joanne Meyer
 ; TITLE OF INVENTION: Detection of Polymorphisms in the Human
 ; FILE REFERENCE: MRI-021
 ; CURRENT APPLICATION NUMBER: US/10/071, 411A
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: 60/267,515
 ; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/314,248
 ; PRIOR FILING DATE: 2001-08-21
 ; NUMBER OF SEQ ID NOS: 66
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 168273
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(168273)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-10-071-411A-2

Query Match 12.3%; Score 1058.6; DB 4; Length 168273;
 Best Local Similarity 73.4%; Pred. No. 9.8e-259; Indels 39; Gaps 12;
 Matches 1526; Conservative 0; Mismatches 514;

Qy 1 TGTATGAAGCCATGTACCTTATATACCAAAACGAGAAAGATATA--CAAAAAGAAA 58
 Db 159686 TTTATAGGCCAGCANTCATCTGTATACAAAGCCGGGAGAGACAAAACAAAAGAG 159745
 Qy 59 ACTATAGACAGTACCATGATGATATATACATGACAGAAATCCCAACAAAATCTAGCTA 118
 Db 159746 ATTTTAGACCAATATCTTGAGGAAACATTGATGCCAAATCTCAATAAATCTGGCAA 159805
 Qy 119 ACCCAATCCACAGCATATCAAGAAATATATCCACCATGTCAGGGGTTTCATACCAAG 178
 Db 159806 ACCAAATCCAGCAGCATATCAAAAGCTTATCCACCATGATCAAGGCTTCATCCCTA 159865
 Qy 179 GGGTGC-AGGATAGGTAACTATACACAGAGTCAATTAATGTATATACATCATAAACAGAA 237
 Db 159866 GGATCCAGAGCTGTTCACATATGCAAAATCAATTAATGTAATCCAGCATATAAACAGAA 159925
 Qy 238 TTTAAAACAAAATCATGATATCTCAATAGATGCTGAAAAAGCATTTGGCAAAATCT 237
 Db 159926 CCAAAAGACAAAACCAATGATATCTCAATAGATGAGAAAAAGGCTTTGACAAAATTC 159985
 Qy 298 AACATTTCTTTATGATTAAACCTTCAGCAAAATTCGACATAGAAAGGACATACCTTAATG 357
 Db 159986 AACACACCTTCATGTTAAACCTCTCAATTAATTAATGATGCTGGGACATATCTCAAA 160045
 Qy 358 TAATAAAAACCATATATGACGAGCCAGCAAGCAAACTTATATCTGAATGGGAAAAGTTGA 417
 Db 160046 TAATAGAGCTATCTATATGCAAAACCAAGCCAAATTCATATGANTGGCAAAAACCTGG 160105
 Qy 418 AAACATTTGCTCCTGAGAACTGGAACAAAGACAGAGATG-CTACTTTTCCACTCTTATCA 476
 Db 160106 AAGCATTCCTCTTGAATAATTTGGCAAAAGAGGATGCCCTCTCTCACTCTCTATTTCA 160165
 Qy 477 ACATAGTAGTGAAGTTTATAGCCAGAGCATCAGACAGAGAAAGAAATCAAGGGCACCC 536
 Db 160166 ACATAGTCTGGAAGATCTGTCCAGGGCAATCAGGAGAGAGAAATTAAGGGTATTC 160225
 Qy 537 AAATCAATTAAGAGAAAGTCAAACTGTCTCTGTTCACTGATGATGATGATTTACCTAG 596
 Db 160226 AATTGGAAGAAAGAGAAAGTCAAAATGTCTCTGTTTGAAGATGACATATTTTATATCTAG 160285
 Qy 597 AAAACCTTAAAGCTATCTCAGAAAGTCTCTAGAACCTGATACATTAATTCAGTAAAGTTT 656
 Db 160286 AAAACCTTAAAGCTATCTCAGAAAGTCTCTCTTAAGCTGATACCACTTCAGCAAAAGTCT 160345
 Qy 657 CAGGATACAAACTAATGTATCAAAATCAGTACAGTGTATACCAACAGTGAACCAAG 716
 Db 160346 CAGGATACAAACTAATGTATCAAAATCAGTACAGTGTATCTTATACCAACAGTGAACCA 160405
 Qy 717 CTGAGATCAAAATCAAGAACTCAAACTTTTACAAATAGCTGTAAA-----AAAATACT 770
 Db 160406 CAGAGAGCAAAATCAATGATGAACTCCCATTCACAAATGCTTCAAGAGAAATTAATATCC 160465
 Qy 771 TTAGAATATTTCTTACCAAGAGAGTGAAGAGCTTACAAAGAAACTTACAAAACACAGC 830
 Db 160466 TAGGAATCACCTTACAAAGGAGTGAAGAGCACTTTCAAAGAGAACTTACAAAACCACTGC 160525

Qy 831 T-----GACATCATATGATGACACAAACAGTGGAAAACATCCCATGCTCATGATGGGTA 886
 Db 160526 TCACGAAATTAAGAGGATACAAACAAATGGAAGACCTTCACTGCTCATGATGGTAA 160585
 Qy 887 GAATCAATATTTGTAATGATGACCATATTTGCCAAAAGCAATCTACAGATTCAATGCAATTC 946
 Db 160586 GAATCAATATTTGTAATGATGACCATATTTGCCAAAAGGTAATTTACAGATTCAATGCAATTC 160645
 Qy 947 CCACCAAAATATCATCATATCTTTCAGGAAGTGAAGAAACAAATCTTAATTAATCTA 1006
 Db 160646 CCATCAAGCTACCAATTTACTTCTTCACTAA--TTGGAAGAAACATCTTTAAAGTTCA 160704
 Qy 1007 TGGACAAACAAACCAAA 1066
 Db 160705 TGGAC-----CAAAAAAAAAAGCCGATTCACAAAGTCAATCTGAGACCAA 160750
 Qy 1067 AGAACAAATCTGAGGACATCACATTAACCATTTTCAAACTATATCTACAAAGCTATATCA 1126
 Db 160751 AGAACAAAGCTGAGGACATCACATCTGACTTCAAGCTATATCTACAAAGCTATATCA 160810
 Qy 1127 CCAAAACATCATGAGCATGACATTAATACTAGGCACTAGACCAATGGAAGAAAGAGAGA 1186
 Db 160811 CCAAAACAGCTGTGTCTGTATCCAAACAGAGATATGATCAATGGAACAGACAGAGC 160870
 Qy 1187 ATCCAGAAATTAAGCCAAATTAATATATAGCCAACTGATTTTGAACAAGCAAAACAA 1246
 Db 160871 CCTCAAAATTAATATGCTGATATGACACATGCTGATCTTTACAAACCTGAGAAACAA 160930
 Qy 1247 TAAAGTGGGAAAAGACATTTCTAGTTAACAAATGCTGTGAGATTTTGGCAAGCCACAT 1306
 Db 160931 AGAAATGGGAAAAGATATCCCTATTTAATAATGAGTGGGAAAACCTGGCTAGCCATAT 160990
 Qy 1307 GTGGAAGATGAACTGATCCCTGTCTCTCACTTAATACAAATGATGATCAAGATGG 1366
 Db 160991 GTGGAAGATGAACTGATCCCTT-CTTACACCTTATACAAATTAATTCAGATGG 161049
 Qy 1367 ATCAAGACTTAAATCTGAGACCTTAAACCAATTAATAATCTGAGATTAATCATGAAAA 1426
 Db 161050 ATTAAGACTTAAATCTGAGACCTTAAACCAATTAATAATCTGAGATTAATCATGAAAA 161109
 Qy 1427 ATGCTTCTAGACATTCATTAAGCAAAAGCTTCATGGCCAGAACCCCAAGTAATGCA 1486
 Db 161110 ACCATTCAGAGACATTAAGCAATGGCAAGGACTTCATGCTTAAACCAAAACCAATGCA 161169
 Qy 1487 ACAAAACCAAAATTAATATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 161170 ACAAAACCAAAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTA 161229
 Qy 1547 ACAATCATTAAGCAGACAAACAGACAAACCCAGAGTGAAGAAATCTTCACAAACTTA 1606
 Db 161230 GAAACTACATCAGAGTGAATGAGCAACTTACAGATGAGAGAAATTTTTCGAAATCTAC 161289
 Qy 1607 GCATCTGACTAAGACTAATATTCGGATTCACAGAGAACTCAAACTAATGCAAGAG 1666
 Db 161290 TCATCTGACAAAGGCTTAATATTCAGAACTTCAAAAGAACTTAACTAATGTTCAAGAA 161349
 Qy 1667 AAAGCAAACTATCCCTGAAAGAGTGGGCTTAAGACATGAATAGCAATTTCTCAAAAGAA 1726
 Db 161350 A-----AAACAAACCCCTATCAAAAGTGGGCAAAATCTGAAAGAAACACTTTCAAAAGAA 161405
 Qy 1727 GATATCAAAATGGCCAAAC--AAACAGAGAAAAATGCTTAACTATCACTAATGATTAAGGAA 1784
 Db 161406 GATATTAAGCAACCAAGACATGAAATGCTATCTATCTATCTATGAGTCAAGAGAA 161465
 Qy 1785 ATGTAATCAACCTGTAAATGCGATACACCTTACTCTGCAAGAAATGCTATTAATTTAA 1844
 Db 161466 ATGCAATCAAAACCAATGATGATGATCATCTCAACCGTATGATGATGATGATGATGAT 161523
 Qy 1845 AATCTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 Db 161524 AATGTCAGAAACAAACAGATGCTGAGAGAGACATGAGAAATGATGATGATGATGATGATGAT 161583

QY	1905	GGGTGGGAATGTTAACTTTCGCAACACATATGAAAAACGTGTGAAATTTCTTAAGAA	1964
Db	161584	TGTTGGGGGCTGTAAATTCAATTCAATCTGTTGMAACAAGTGTGGCAATTCCTCAAGGAT	161643
QY	1965	CTAAAGTAGATCGACCATTTTGATCCAGCAATCCCATTAATATGTTAATAATATATAT	2024
Db	161664	CTAGAACTAGAAATATCCATTGACCCAGCAATCCATT-ACGTGGGTATATACCGAAAAGA	161702
QY	2025	TTATATACCATGGAATATACAATCAGCCATTAATAAAAAAAT	2063
Db	161703	TTACAAATGTGTCAACTATAAAGACATGACACGAAT	161741

RESULT 15
US-09-751-389-3/c
; Sequence 3, Application US/09751389
Date Rec'd: 05/03/04

```

? TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
? TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
? TITLE OF INVENTION: THEREOF
? FILE REFERENCE: C0001067
? CURRENT APPLICATION NUMBER: US/09/751,389
? CURRENT FILING DATE: 2001-01-02
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 786431
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc.feature
? LOCATION: (1)...(786431)
? OTHER INFORMATION: n = A,T,C or G
? US-09-751-389-3

```

Query Match	12.2%	Score 1054.4	DB 4	Length 786431
Best Local Similarity	73.9%	Pred. No. 3.1e-257		
Matches 1510, Conservative	0	Mismatches 496	Indels 38	Gaps 12

QY	9	SCCATGTCACCTTTAAATCCAAAAACGAGGAAAGGATAT--CAAAAAAGAAACTATGA	66
Db	579365	GCACGATCATCTCTGATACCAAGCTGGCAGAGACACAAAAAAGGAATTTTGA	579306
QY	67	CCAGTACCACGTAGATATATACATGCAGAAATCCCAAAAAATACGTAAACCAATC	126
Db	579305	CCAATGTCCTTATGAAATCATGCATGCAAAAAATCTCAGTAAATACTGGCAAAACCGATC	579246
QY	127	CAACAGCATTTCAAGAAGTATATCCACCATTGTTCAGATGGGTTTCATCCAGGGGTGC	A 185
Db	579245	CAGCAGCACATCAGAAAGTTATCCACCATGATCAAGTGGGCTTCATCCGTGGAGATGCA	579186
QY	186	GGATAGTTTACATACACAGTCAATTAATGTGATCATCACATAACAGAAATTAAAAAC	245
Db	579185	GGCTGGTTCAACATACGGAAATCAATTAATGTATATCAGCATTTAAACGAAACCAAGAC	579126
QY	246	AAAAATCAGTATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTAACATTTTC	305
Db	579125	AAAAACCAATGATTTATCTCAATAGATGAGAAAGCGCTTTGACAAAAATTCACCAACCC	579066
QY	306	TTTATGATTTAAAACTTCGACGAAATCGACATAGAAAGGACATACCTTAATGATATAAA	365
Db	579065	TTCAATGCTAAAAACCTCTCAATTAATTTAGTATTTGATGGGACGTATCTCAAAATATATAGA	579006
QY	366	GCATATATGACGAGCCACACAGCAAACTATTATCTGAATGGGGGAAAAAGTTGAAAAATTC	425
Db	579005	GCTATCTATGACAAACCCACACCAATATCATACTGAATGGGCAAAAACTGGAAGCATTC	578946
QY	426	TCCCTGAGAACTGGAAACAAAGACAGATG--CTACTTTCACCTTCTAATTTCAACATATGA	484
Db	578945	CTTTGAAAAATGGGCAACAGACAGGAGATCCTTCTCTCACCACTCTCATTAATATATATG	578866

QY	485	TTGGAAGTTTATGACGAGGCAATCAGACAAGGAAAGAAATCAAGGGGACCCAAATCAAT	544
Db	578885	TTGGAACTTCTGGCTTAGGACAAATCAACGACGAGGAGGAAATTAAGGGATTTCAATTAGA	578828
QY	545	AAAGAGAAAGTCAAACTGTCCCTGTTCACTGATGTATGATTTGATACCTAGAAAACCT	604
Db	578825	AAAGGAGAAAGTCAAAATTTGTCCTGTTGGCAGATGACATGATTTGTATATCCAGAAAACCC	578766
QY	605	AAAGCTCATCCAGAAAGCTCCTGAACTGATATCAATTAATTCAGTAAAGTTTCAGGATAC	664
Db	578765	ATCATCTGAGCCCAAAATCTCCACAAGCTGATTAACCACTTCAGCAAAAGTCTCAGGATAC	578706
QY	665	AAACTTAATGTATCAAAATTCAGTAGCATCTGCTATATACCAACAGTGAACCACTGAGAA	724
Db	578705	AAAACTAAATGTACAAAATAATCAACACATTTCTTATACCCAA---TAAAGAACAAGACC	578650
QY	725	CAAAATCAAGAATCAAACTTTTCAATAGCTGTAAA-----AAATATCTTAAGAAAT	778
Db	578649	CAAAATCAAGAATCAAACTTTTCAATAGCTGTAAA-----AAATATCTTAAGAAAT	578590
QY	779	TTCTTACCAGAGAGTGAAGACCTCTTACAAAGAAAATCAAAACACACGCT---GAC	834
Db	578589	CAACTCAAAAGGATGTGAAGACCTCTTCAAGAGAACTAACAAACACTGCTCAAGAA	578530
QY	835	ATCATATGATGACAAACAAGTGAACAACATCCCATCTCATGATGGTATGAATCAAT	894
Db	578529	ATTAAGAAGGAGACAAACAAATGTGAAGAACATTCATCTCATGGGTAGGAAGATCAAT	578470
QY	895	ATTGGAATAATGACCATATGTCGCAAAAGCAATCTCAAGTTCAATGCAATTTCCACAA	954
Db	578469	ATCTGAAAATATGGCATATCTGCCCAAGTAATTTATATGATTCATATGCCATCCCCATCAAG	578410
QY	955	ATATCATCATGATTTCTTACAGAACTAGAAAAAAACAATTCATAATTCATATGACAACA	1014
Db	578409	CTACCAAGCATTTCTTACAGAA---TTGAAAAAACTACTTAAAGTTCATATGGAAC--	578353
QY	1015	CAACCAAAAAAAAAAAAAAAAACCCGCATGCGCAAGCAAGCACTTAGCAAAAAAGACAA	1074
Db	578352	-----CAAAAAAGCCCACTAGGCAAGTCAATCTTAAGCCAAAAAGACAA	578305
QY	1075	TCTGGAGGCATCACTTACCATCTTCAACATATCTCAAGGGATATATCAACCAAAACA	1134
Db	578304	GCTGAGGCGACACGCTTACTGACTTCAAACTATATCTCAAGGCTTACGTAAACCAAAACA	578245
QY	1135	TCATGCGCATGACATTAATACTAGGCACTAGACCAATGGAAGAAAGAGAGATCCAGAA	1194
Db	578244	GCAATGTATCTGTATCCAAACAGAGATATAGCAATGGAACAGAACAGAGCCCTCAGAA	578185
QY	1195	ATTAAGCCAAATTAATATAGCCCACTGATTTTGACAAAAGCAACAAACATTAAGTGG	1254
Db	578184	ATTAATGTCTGACATATCTACAGCTATCTGTATCTTGAACAACTGAAACAAACAGCATGG	578125
QY	1255	GGAAGAAAGCATTTAGTTAAACAAATGAGTGTGAGATTATGCGCAAGCCACATGTGGAAGA	1314
Db	578124	GGAAGAAAGTTCCTTATTTTAATTAATGTGTGGGAAAACTGGCTAGCCATATGTGAAAG	578065
QY	1315	ATGAAACTGATCCCTGTCTCTCACTTAATACAAAAATGATACAAAGTGTATCAAGA	1374
Db	578064	CTGAAACTGATCCCTTC-CTTAACCTTATATCAAAAATTCATTCAGAATGATTTAAAG	578006
QY	1375	CTTAATTTGAGACTTAACCAATTAATAATTTCTAGAGATTAATCATCAGAAAAATGCTTCT	1434
Db	578005	CTTAATTTGATGACTTAACCAATTAATAATTTCTAGAGATTAATCATCAGAAAAATGCTTCT	577946
QY	1435	AGACATTCATTAAGCAAGACCTTCAAGGCAAGAACCCAAAGTAATGTCAACAAAAAC	1494
Db	577945	GGAATTAAGCAATGGGCAAGACCTTCAATGTCTTAACACCAAAAG-CATGGCAACAAAGA	577887
QY	1495	AAAAATTAATGATAGACTTAATTTAACTAAAAAGCTTTTGGCAGCAAAAAACATCAT	1554
Db	577886	CAAAATTAAGCAAAATGGGATCTTAGTTAACTAAAGAGCTTTGCAACAGCAAAAGAAACATAC	577827
QY	1555	TAGCAGAGCAAAACAGACCAACCGAGTGAAGAAAAATCTTACAAACTAAGACTTGA	1614

QY	1015	CAACCAAAAAAAAAAACC	CGCATAGCCAAAGCAAGCTTAGCAAAAAAGAACAA	1074
Db	578352	-----CAAAAAAG	CCCAATAGCAAGTCAATCTTAAGCCAAAAAGAACAA	5783050
QY	1075	TCGTGGAGGCATCACA	CTTACCATCTTCCAACTATCTCAAGGCGTATATATCAACCAAAAAA	1134
Db	578304	GCTGGAGGCACACG	CTACCTCGACTTCAAACTATATCTCAAGGCTACAGTAAACCAAAACA	578245
QY	1135	TCATGGCACTGACAT	TAAACTAGGCACATAGCACAATGGAAAAAGAGAGAAATCCAGAA	1194
Db	578244	GCATGTATCTGTAT	CCAAACAGAGATATAGCCATAGGAAACAGAACAGAGCCCTCAGAA	578185S
QY	1195	ATAAAGCCAAATTA	ATTATAGCCAATGATTTTTGCAAAAGCAACAAAAAACTAAAGTGG	1254
Db	578184	ATAATGCTGCATAT	CTACAGCTATGTATCTTTGCAAAACCTGAGAAAAACAAGCAATGG	578125S
QY	1255	GGAAAGACATTTAG	TTAACAAATGTGTGAGATATTGGCAAGCCACATGTGGAGA	1314
Db	578124	GGAAAGAGATTCCT	TAATTAATAAGTGTGGGAAAACTGGCTAGCCATATGTAGAAAG	578065S
QY	1315	ATGAAACTGGAAT	TCCTTGTCTCACTTAATACAAAAATTGATACAAAGTGAATCAAAAGA	1374
Db	578064	CTGAAACTGGAAT	CCCTTC - CTTACACCTTAATACAAAAATCAATTCAAAGTGAATTAAMAG	578006
QY	1375	CTTAAATCTGAG	ACCTTAACCATTAATAATCTAGAACATACACAGAAAAATGCTTCT	1434
Db	578005	CTTAAATGTTGA	CCCTTAACCATTAATAAACCTTAGAAGAAAACTTAGGCATATCAATTA	577946
QY	1435	AGACATTCATTA	GCCAAAGACTCATGGCCAAAGAACCCAAAGTAAATGCAACAAAAAC	1494
Db	577945	GGACATTAAGCAT	TGGCCAAAGACTTCATCTTAAAAACACCAAAAG - CATGGCAACAAAABA	577887
QY	1495	AAAAATTAAT	TAGACTTAATTTAAACTAATAAGCTTTTGGCAGCAAAAAACAATCAT	1554
Db	577886	CAAAATTTGAC	AAATGGGATCTTAGTTAAACTTAAGAGCTTTGCAACAGCAAAAGAAACTAC	577827
QY	1555	TAGCAGAGCAAA	CAGACCAACCCAGATGAGAGAAAAATCTTACAAACTAAGCATCTGA	1614

Db	577826	CATTAGACTGAAACAGGCACTCAACAAGTGGGAGAAAATTTTGCATCTATTCATCTGA	57767
QY	1615	CTAAGGACTAATATATCCGAAATCCACAGGAATCTCAACAAATCAGACAGAGAAAGCAA	1674
Db	577766	CAAAAGGCTAAATATCCAGATCTACAAATGAACTCAACAAATTTACAGAAAAATCAAA	577707
QY	1675	CAATCCCATGAAGAAGTGGGCTAAGGACATGAATTAGACAATCTCAAAAGAAATATACA	1734
Db	577706	CAACCCCATCAAAAAGTGGGCAAAAGATATGAACAGACACTTCTCACAGAGACATTTA	577647
QY	1735	AATGGCAACAA-ACAGGAAAAAATGCTTAACATGCTAATGATTAGGAAATGTAAAT	1792
Db	577646	TGCAGCCAAAAACACAGAAAAAATGCTCATCATCTGGCCATCAGAGAAATGCAAT	577587
QY	1793	CAACCTGTAATGCGATACCACTTACTCTGCAAGATGGTCATTAATTTAAAAATCTAA	1852
Db	577586	CAAAACCAATGGGATACCATCTCACACCAATTAGATGGCGATCATTAAMAAATCAGG	577527
QY	1853	AAATAATAGATGTGGTGGGTCTGTGTGATTAAGAAACACTTTACACTGCTGGTGGGA	1912
Db	577526	AAACACAGGTCTGGAGAGATGTGAGAAATAGAAATACCTTTTCACTGTTGTGGGA	577467
QY	1913	ATGTAACCTTGGCGAACCACTATGAAAAAGTGTGAAATTTCTTAAGGAACCTAAAGT	1972
Db	577466	CTGTAACTAGTTCAACCAATGTGGAAGTCAGTGTGTGATTTCTCAGGATCTAGAACT	577407
QY	1973	AGATGACCATTTGATCCAGCAATCCCAATTAATATATATTAATATATATATATAC	2032
Db	577406	AGAAATACCATTTGACCCAGATCCCATTT-ACTGGGTATATACCCAAAGATTAATAAT	577348
QY	2033	CATG 2036	
Db	577347	CATG 577344	

Search completed: December 2, 2004, 18:17:25
 Job time : 730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 01:15:24 ; Search time 4225 Seconds
(without alignments)
11213.321 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622
Sequence: 1 tgcataagccaatgcacac.....gtccccaagcccttacc 8622

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3694831 seqs, 2747406616 residues

Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8622	100.0	8622	11	US-09-867-570-3
2	1429.4	16.6	398287	17	US-10-741-601-5719
3	1401	16.2	66494	17	US-10-450-826-47
4	1395	16.2	209484	13	US-10-087-192-418
5	1395	16.2	209484	13	US-10-331-053-4
6	1390.4	16.1	465237	9	US-09-933-267A-1
7	1375.6	16.0	1691139	14	US-10-067-514-1
8	1375.6	15.9	1691139	16	US-10-419-723-1
9	1374.6	15.9	1400	14	US-10-183-116-30
10	1374.6	15.9	1400	15	US-10-225-567A-673
11	1370	15.9	100944	17	US-10-322-696-4
12	1366.4	15.8	277616	17	US-10-367-094-83

C	13	1358	15.8	99014	9	US-09-880-107-3428	Sequence 3428, Ap
	14	1346	15.6	136284	17	US-10-775-169-148	Sequence 149, Ap
	15	1311.6	15.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
	16	1311.6	15.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
	17	1311.6	15.2	1503841	9	US-09-946-807-1	Sequence 1095, Ap
	18	1308	15.2	358246	15	US-10-292-798-1095	Sequence 2032, Ap
	19	1302.6	15.1	81098	13	US-10-087-192-2032	Sequence 7, Appl1
	20	1302	15.1	145608	17	US-10-721-693-7	Sequence 7, Appl1
	21	1302	15.1	145606	18	US-10-852-997-7	Sequence 8197, Ap
	22	1301.8	15.1	32167	10	US-09-764-891-8197	Sequence 77, Appl1
	23	1289	15.0	290547	17	US-10-367-094-77	Sequence 28, Appl1
	24	1288.6	14.9	247682	16	US-10-235-192A-28	Sequence 3, Appl1
	25	1286.8	14.9	148567	9	US-09-801-8768-3	Sequence 3, Appl1
	26	1286.8	14.9	148567	14	US-10-254-869-3	Sequence 3, Appl1
	27	1286.8	14.9	148567	16	US-10-667-442-3	Sequence 166, Appl
	28	1286.8	14.9	59475	17	US-10-322-696-166	Sequence 1273, Ap
	29	1274	14.8	1369	15	US-10-292-798-1273	Sequence 8024, Ap
	30	1271.6	14.7	32249	10	US-09-764-891-8024	Sequence 32, Appl
	31	1264	14.7	1604	14	US-10-183-116-32	Sequence 688, Appl
	32	1264	14.7	1604	15	US-10-225-567A-688	Sequence 58, Appl
	33	1258.2	14.6	246940	17	US-10-322-696-58	Sequence 2002, Ap
	34	1247.8	14.5	261817	13	US-10-087-192-2002	Sequence 5611, Ap
	35	1247.4	14.5	253861	17	US-10-741-601-5611	Sequence 40, Appl
	36	1247	14.5	79684	15	US-10-034-650-40	Sequence 149, Appl
	37	1245.6	14.4	136284	17	US-10-775-169-149	Sequence 5719, Appl
	38	1234.8	14.3	398287	17	US-10-741-601-5719	Sequence 15, Appl1
	39	1228.4	14.2	2040	14	US-10-183-116-15	Sequence 427, Appl
	40	1225.8	14.2	267156	11	US-09-968-007A-427	Sequence 154, Appl
	41	1225.8	14.2	684973	9	US-09-263-959-1	Sequence 28, Appl
	42	1219.8	14.1	49536	13	US-10-087-192-154	Sequence 1059, Ap
	43	1219.8	14.1	49536	18	US-10-331-053-28	Sequence 10, Appl
	44	1208.8	14.0	21188	15	US-10-292-798-1059	
	45	1203	14.0	130505	17	US-10-367-094-10	

ALIGNMENTS

RESULT 1
US-09-867-570-3
Sequence 3, Application US/09867570
Publication No. US20040076951A1
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1009090-CIP
CURRENT APPLICATION NUMBER: US/09/867,570
CURRENT FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 09/695,045
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 8622
TYPE: DNA
ORGANISM: Human
US-09-867-570-3

Query Match 100.0%; Score 8622; DB 11; Length 8622;
Best local similarity 100.0%; Pred. No. 0;
Matches 8622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TGATGAGCGCATGCTTATATACCAACAGGAAGATATCAAAAAGAAAC	60
Db	1	TGATGAGCGCATGCTTATATACCAACAGGAAGATATCAAAAAGAAAC	60
Qy	61	TATGACCACTACCTGATGATATATACGAGAAATCCCAAAATATAGCTTAC	120
Db	61	TATGACCACTACCTGATGATATATACGAGAAATCCCAAAATATAGCTTAC	120
Qy	121	CCATCCACAGCATATCAAGAAATATCCACCATTTGTCAAGTGGTTTCATACAGG	180

Db	121	CCAAATCCAAACGCAATATCAAGAAAGTAATCCAACTGTCAAGTGGTTTCAATACCGAGG	180
Oy	181	GTGCGAGTAGGTTAAACATACCAAGTCAATTAATGTATACATCACTAAACAGAAATTA	240
Db	181	GTGCGAGTAGGTTAAACATACCAAGTCAATTAATGTATACATCACTAAACAGAAATTA	240
Oy	241	AAAACAAAATCACATGTATCATCTCAATAGATGCTGAAAAAGCATTTTGACAACAAATCTAAC	300
Db	241	AAAACAAAATCACATGTATCATCTCAATAGATGCTGAAAAAGCATTTTGACAACAAATCTAAC	300
Oy	301	ATTTCTTTATGATTTAAACCTTCACCAAAATTCGAAATGAAAGGACATACCTTAAATGTAA	360
Db	301	ATTTCTTTATGATTTAAACCTTCACCAAAATTCGAAATGAAAGGACATACCTTAAATGTAA	360
Oy	361	TAAAGCCATATATGACGGAACCCACAGCAAAACATATATCTGAATGSGGAAAAAGTTGAAA	420
Db	361	TAAAGCCATATATGACGGAACCCACAGCAAAACATATATCTGAATGSGGAAAAAGTTGAAA	420
Oy	421	CATTGTCCCTGAGAACTGGAACAGAACAGAGTGTACTTTTCAACACTTATTTCAACAT	480
Db	421	CATTGTCCCTGAGAACTGGAACAGAACAGAGTGTACTTTTCAACACTTATTTCAACAT	480
Oy	481	AGTAGTGAAGTTTATAGCCAGAGCAATCAGACAAAGAAAGAAATTCAGAGGGCACCCCAAT	540
Db	481	AGTAGTGAAGTTTATAGCCAGAGCAATCAGACAAAGAAAGAAATTCAGAGGGCACCCCAAT	540
Oy	541	CAATTAAGAGNAGTCAAACTGTCCCTGTTCACTGATATATGATTTGATATCTAGAAAA	600
Db	541	CAATTAAGAGNAGTCAAACTGTCCCTGTTCACTGATATATGATTTGATATCTAGAAAA	600
Oy	601	CCCTTAAGAATCATATCCAGAAAGCTCTTGAAACTGATATACATAAATTCAGTAAAGTTTCAGG	660
Db	601	CCCTTAAGAATCATATCCAGAAAGCTCTTGAAACTGATATACATAAATTCAGTAAAGTTTCAGG	660
Oy	661	ATACAAACTTAATGTACACAAATTCAGTAGCACTGCTATATACCAACAGTGAACAAAGCTGA	720
Db	661	ATACAAACTTAATGTACACAAATTCAGTAGCACTGCTATATACCAACAGTGAACAAAGCTGA	720
Oy	721	GAATCAAAATCAAGAACTCAAAACACTTTTACAAATAGCTGTAAAAAAAATCTTAAGAATATT	780
Db	721	GAATCAAAATCAAGAACTCAAAACACTTTTACAAATAGCTGTAAAAAAAATCTTAAGAATATT	780
Oy	781	CTTACCCAGAGAGTGAAAGACCTTCTACAGAGAAAACTACAAAACACAGCTGACATCATTA	840
Db	781	CTTACCCAGAGAGTGAAAGACCTTCTACAGAGAAAACTACAAAACACAGCTGACATCATTA	840
Oy	841	GATGACACAAACAAAGTGGAAACACATCCCATCTGATGAGGTGGATGAGATCAATATTGTG	900
Db	841	GATGACACAAACAAAGTGGAAACACATCCCATCTGATGAGGTGGATGAGATCAATATTGTG	900
Oy	901	AAAAAGACAAATTTTSCAAAAAGCAATCTACAAAGTTCAATGTCAATCCCAACCAAAATATCA	960
Db	901	AAAAAGACAAATTTTSCAAAAAGCAATCTACAAAGTTCAATGTCAATCCCAACCAAAATATCA	960
Oy	961	TCATCATTTCTTCACAGAACTAGAAAAAAAACAATTTCTAAAAATTCATATGGAACAACACA	1020
Db	961	TCATCATTTCTTCACAGAACTAGAAAAAAAACAATTTCTAAAAATTCATATGGAACAACACA	1020
Oy	1021	AAAAAAAACAAAAAACCCTGCTAGCCAAAGCAAGCTTAGCAAAAAGAAACAAATCTGGA	1080
Db	1021	AAAAAAAACAAAAAACCCTGCTAGCCAAAGCAAGCTTAGCAAAAAGAAACAAATCTGGA	1080
Oy	1081	GGCATCAACATTAACCATCTTCAAACTATCTACAAAGGCTATATATACCAAAAACATCAATG	1140
Db	1081	GGCATCAACATTAACCATCTTCAAACTATCTACAAAGGCTATATATACCAAAAACATCAATG	1140
Oy	1141	CACCTGACATTAATATAGCCAACTGATTTTGTACAAAGCAACAAAACATTAAGTGGGAAAA	1200
Db	1141	CACCTGACATTAATATAGCCAACTGATTTTGTACAAAGCAACAAAACATTAAGTGGGAAAA	1200
Oy	1201	CCAAATTAATATATAGCCAACTGATTTTGTACAAAGCAACAAAACATTAAGTGGGAAAA	1260

Db	1201	CCAAATATATTATAGCCAACTGATTTTGGACAAACCAACAAAACATTAAGTGGGAAAA	1260
Qy	1261	GACATCTTAGTAAACAAATGSGTGTGAGATTATTTGGCAAGCCACATGTGGAGATGAA	1320
Db	1261	GACATCTTAGTAAACAAATGSGTGTGAGATTATTTGGCAAGCCACATGTGGAGATGAA	1320
Qy	1331	CTGGAATCCCTTGTCTCTCACTTAATATACAAAATTGATATACAAATGATTCAAAGCTTAA	1380
Db	1331	CTGGAATCCCTTGTCTCTCACTTAATATACAAAATTGATATACAAAGATGATTCAAAGCTTAA	1380
Qy	1381	TCTGAGACCTTAAACCATTAATAATTCTGAGAGATATACATCAGAAAATATGCTTCTGACAT	1440
Db	1381	TCTGAGACCTTAAACCATTAATAATTCTGAGAGATATACATCAGAAAATATGCTTCTGACAT	1440
Qy	1441	TCACCTTAGGCAAAAGACTTCATGGCCAGAGAACCCAAAAGTAAATSCAACAAAACAAAAT	1500
Db	1441	TCACCTTAGGCAAAAGACTTCATGGCCAGAGAACCCAAAAGTAAATSCAACAAAACAAAAT	1500
Qy	1501	AAATATGATAGACTTAATTAATCTTAATAAGCTTTTGGCAGACAAAACATCATTTACAG	1560
Db	1501	AAATATGATAGACTTAATTAATCTTAATAAGCTTTTGGCAGACAAAACATCATTTACAG	1560
Qy	1561	AGCAAAAGAGAACCCACCGAGTGGAGAAAATCTTCACAACTATGATCTGACTAGG	1620
Db	1561	AGCAAAAGAGAACCCACCGAGTGGAGAAAATCTTCACAACTATGATCTGACTAGG	1620
Qy	1621	ACTATATATCCGGAATCCACAAGGAATCTCAACAAATCAGACAGAGAAAGCAACATCC	1680
Db	1621	ACTATATATCCGGAATCCACAAGGAATCTCAACAAATCAGACAGAGAAAGCAACATCC	1680
Qy	1681	CATGAAAAGATGGGCTAAAGGACATGATATGACATTTCTCAAAAGAAATATACAAATGGC	1740
Db	1681	CATGAAAAGATGGGCTAAAGGACATGATATGACATTTCTCAAAAGAAATATACAAATGGC	1740
Qy	1741	CAACAAAAGGAAAAAATGCTTAAACATCATAATGATAGGGAATGTAATCAACTG	1800
Db	1741	CAACAAAAGGAAAAAATGCTTAAACATCATAATGATAGGGAATGTAATCAACTG	1800
Qy	1801	TAATGCGATACCACTTACTCTCGCAGAGATGTCATTAATTTAAATACTTAATAATA	1860
Db	1801	TAATGCGATACCACTTACTCTCGCAGAGATGTCATTAATTTAAATACTTAATAATA	1860
Qy	1861	GATGTTGGTGGTCTGTGTGTGATTAAGGACACTTTTACATGCTGGTGGGAATGTAAAC	1920
Db	1861	GATGTTGGTGGTCTGTGTGTGATTAAGGACACTTTTACATGCTGGTGGGAATGTAAAC	1920
Qy	1921	TTGGSCAACACTATNGGAAAACAGTGTGGAAATTTCTTAAGAACCTAAAGTAAATGAC	1980
Db	1921	TTGGSCAACACTATNGGAAAACAGTGTGGAAATTTCTTAAGAACCTAAAGTAAATGAC	1980
Qy	1981	CATTGATATCCAGCAATCCCATTTAATATGATTAATATATATATTTATATATCATGCAAT	2040
Db	1981	CATTGATATCCAGCAATCCCATTTAATATGATTAATATATATATTTATATATCATGCAAT	2040
Qy	2041	ACAACTCAGCCATTAATAAAGATTAATATGATGACATTCACGCAATCTAGATGAAATGG	2100
Db	2041	ACAACTCAGCCATTAATAAAGATTAATATGATGACATTCACGCAATCTAGATGAAATGG	2100
Qy	2101	AGACCCCTATTTCTAAGTGGGGTAACTCAGGAATGGAAAAACCAACATCATATGTTCTCAC	2160
Db	2101	AGACCCCTATTTCTAAGTGGGGTAACTCAGGAATGGAAAAACCAACATCATATGTTCTCAC	2160
Qy	2161	TTACAAAGTGGGGCTTAAGCTGTGAGAGACAGCAAGGCAATGATATTAATGAATCTGCG	2220
Db	2161	TTACAAAGTGGGGCTTAAGCTGTGAGAGACAGCAAGGCAATGATATTAATGAATCTGCG	2220
Qy	2221	GGACTTGAAGGGAGATGAAAGAGAGGCGAGGAGATTAAGAATCAACAATGSGTACAGT	2280
Db	2221	GGACTTGAAGGGAGATGAAAGAGAGGCGAGGAGATTAAGAATCAACAATGSGTACAGT	2280
Qy	2281	GTAACCTGCTCAGGTGATGGGTGACCAAAATTTCCAGAAATTACCACTTAAGAACTTATC	2340
Db	2281	GTAACCTGCTCAGGTGATGGGTGACCAAAATTTCCAGAAATTACCACTTAAGAACTTATC	2340
Qy	2341	GTAACCTGCTCAGGTGATGGGTGACCAAAATTTCCAGAAATTACCACTTAAGAACTTATC	2400
Db	2341	GTAACCTGCTCAGGTGATGGGTGACCAAAATTTCCAGAAATTACCACTTAAGAACTTATC	2400

QY 2341 CATGGAGCAAAACACACCTGTTCCCAAAATCCCAATGAATAATAATAATAATA 2400
 DB 2341 CATGGAGCAAAACACACCTGTTCCCAAAATCCCAATGAATAATAATAATAATA 2400
 QY 2401 ATGATTTAATTTTCAGAGATTTAAAAAGTTCACGTTCAGAGTTTAAATAAGAGTA 2460
 DB 2401 ATGATTTAATTTTCAGAGATTTAAAAAGTTCACGTTCAGAGTTTAAATAAGAGTA 2460
 QY 2461 AGAATGAAAAGTGTAGCAAGTGTAGCCTGTGCAAAATGGGACTGTAGATTTTCACCTTG 2520
 DB 2461 AGAATGAAAAGTGTAGCAAGTGTAGCCTGTGCAAAATGGGACTGTAGATTTTCACCTTG 2520
 QY 2521 CATACACTTCTGTGCAATTTGGAAGAAAGTATACATGATATATATACCACTATGATA 2580
 DB 2521 CATACACTTCTGTGCAATTTGGAAGAAAGTATACATGATATATATACCACTATGATA 2580
 QY 2581 AAGAAAATCATCAAAAATGTGTGAGGCAATGTGACGCTTGAAATGGTCCCATGATCTAC 2640
 DB 2581 AAGAAAATCATCAAAAATGTGTGAGGCAATGTGACGCTTGAAATGGTCCCATGATCTAC 2640
 QY 2641 TTTTTCATTTGATATTAAGCCTCATATGATGATTCACATGCTTAATGTATGATGCTTAC 2700
 DB 2641 TTTTTCATTTGATATTAAGCCTCATATGATGATTCACATGCTTAATGTATGATGCTTAC 2700
 QY 2701 GCCCATTAATGATTTTAAATCAAGACGCAATTAATTAACAGAGATTGAACAAGATGG 2760
 DB 2701 GCCCATTAATGATTTTAAATCAAGACGCAATTAATTAACAGAGATTGAACAAGATGG 2760
 QY 2761 GACGTATAGAGAGGCTTAAATGTATGATGATGAGGCAAGAGCAAGATTCATCTAC 2820
 DB 2761 GACGTATAGAGAGGCTTAAATGTATGATGATGAGGCAAGAGCAAGATTCATCTAC 2820
 QY 2821 TTAGATTTGTGTCTCAAAATACCTCTGACCTGTATTTGCCATAGTCTCTCATACAGAA 2880
 DB 2821 TTAGATTTGTGTCTCAAAATACCTCTGACCTGTATTTGCCATAGTCTCTCATACAGAA 2880
 QY 2881 ATAAACAAGACTGTACAGATTTTGTGAAGCTGTGATGCTCAACAGCTTTTCAATTCAGCT 2940
 DB 2881 ATAAACAAGACTGTACAGATTTTGTGAAGCTGTGATGCTCAACAGCTTTTCAATTCAGCT 2940
 QY 2941 CCTGTAGGACCTCTCTGAATTAAGCAACAGAAAAGCTCTGAAGTCACTGAATCCCA 3000
 DB 2941 CCTGTAGGACCTCTCTGAATTAAGCAACAGAAAAGCTCTGAAGTCACTGAATCCCA 3000
 QY 3001 GAAAGGCTCTCTACCTTTAGCAAAAGGAGGTCTTTCACACATGCAAAAGAGAGAT 3060
 DB 3001 GAAAGGCTCTCTACCTTTAGCAAAAGGAGGTCTTTCACACATGCAAAAGAGAGAT 3060
 QY 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGCTGATATGATTTTGTCTGAATATC 3120
 DB 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGCTGATATGATTTTGTCTGAATATC 3120
 QY 3121 ATGTCTTATGCTTCCACTTGTGCTACATGACGAGGGGTCAAGAGCTGTTTACAGAAA 3180
 DB 3121 ATGTCTTATGCTTCCACTTGTGCTACATGACGAGGGGTCAAGAGCTGTTTACAGAAA 3180
 QY 3181 AGACAGAGACATGAGCTTCTTTCAGAAAATGAGTGTGTCAACCAAACTGTGTAG 3240
 DB 3181 AGACAGAGACATGAGCTTCTTTCAGAAAATGAGTGTGTCAACCAAACTGTGTAG 3240
 QY 3241 CTCTAAATGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3300
 DB 3241 CTCTAAATGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3300
 QY 3301 TTCAATCTGCTCATGAGAGACTGCTGCTCTTCAATTTCAATTTCAAGAGAGGAGAC 3360
 DB 3301 TTCAATCTGCTCATGAGAGACTGCTGCTCTTCAATTTCAATTTCAAGAGAGGAGAC 3360
 QY 3361 ATGACGACATTTCTTCCAGAGTTGAACCTGCTATAGAGCCAGTTTCTTGTTCACCTTAC 3420
 DB 3361 ATGACGACATTTCTTCCAGAGTTGAACCTGCTATAGAGCCAGTTTCTTGTTCACCTTAC 3420

QY 3421 TTTTCAATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3480
 DB 3421 TTTTCAATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3480
 QY 3481 ATAAATTAAGAGTGTTCATGAGAGAAAGTATTTTCAAGATATGACAGAGTTTAAAGGTAC 3540
 DB 3481 ATAAATTAAGAGTGTTCATGAGAGAAAGTATTTTCAAGATATGACAGAGTTTAAAGGTAC 3540
 QY 3541 GATCTTAAGCAATATCATAGGAGCACTACTTCTCTCCCTTAGGCTGAAAACGACAG 3600
 DB 3541 GATCTTAAGCAATATCATAGGAGCACTACTTCTCTCCCTTAGGCTGAAAACGACAG 3600
 QY 3601 GATCTTAAGCAATATCATAGGAGCACTACTTCTCTCCCTTAGGCTGAAAACGACAG 3660
 DB 3601 GATCTTAAGCAATATCATAGGAGCACTACTTCTCTCCCTTAGGCTGAAAACGACAG 3660
 QY 3661 GAGAGGCAAGTGTGAGAGAGGCTGCTGAGGCAAGGCAAGGCAAGGCTGATATGAT 3720
 DB 3661 GAGAGGCAAGTGTGAGAGAGGCTGCTGAGGCAAGGCAAGGCAAGGCTGATATGAT 3720
 QY 3721 TGATCTGTGTCTCCACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
 DB 3721 TGATCTGTGTCTCCACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
 QY 3781 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 3840
 DB 3781 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 3840
 QY 3841 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 3900
 DB 3841 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 3900
 QY 3901 GCGTGTAGCACT 3960
 DB 3901 GCGTGTAGCACT 3960
 QY 3961 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 4020
 DB 3961 CTTGTGTGAGATATATTAATGATCAAGGAGATGTTTGTGATGATTTTAAACACAT 4020
 QY 4021 TATGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4080
 DB 4021 TATGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4080
 QY 4081 TCCAGTCTGAGATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4140
 DB 4081 TCCAGTCTGAGATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4140
 QY 4141 CCAAGATATGAGAAATCCCAAGGAGTCTTCTGCTGTCTTCAAGTCTCTCTGCTGTCTC 4200
 DB 4141 CCAAGATATGAGAAATCCCAAGGAGTCTTCTGCTGTCTTCAAGTCTCTCTGCTGTCTC 4200
 QY 4201 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCAAGTGTGTATATA 4260
 DB 4201 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCAAGTGTGTATATA 4260
 QY 4261 GAAAGCACTCTCTGAGATGCAAAAGATTAAGAAATGAAAGCAATCTCTCATAG 4320
 DB 4261 GAAAGCACTCTCTGAGATGCAAAAGATTAAGAAATGAAAGCAATCTCTCATAG 4320
 QY 4321 TAAATGAGATATATCCCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4380
 DB 4321 TAAATGAGATATATCCCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4380
 QY 4381 TGAATGATCCCTGTCTATTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4440
 DB 4381 TGAATGATCCCTGTCTATTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4440
 QY 4441 TCTACTGTCACTTTATGAGAGAAATGTTTGCATTTGTTAAAAATGATAGAGAAATTAAT 4500
 DB 4441 TCTACTGTCACTTTATGAGAGAAATGTTTGCATTTGTTAAAAATGATAGAGAAATTAAT 4500
 QY 4501 GTAAATTTTAAAGAAATATATATTTTGTGTTAGAAATTAAGTTTGGCTATCTAATAA 4560

Db 4501 GTAAATTTAAAGAAATATATATTTTGTATGATTAATAGTTTGCTGATCTTAATTA 4560
Qy 4561 GACATGAAGAGAAATATCTTAAACAGAAAGTATAGTTGCTCTGGGTCACTAGTT 4620
Db 4561 GACATGAAGAGAAATATCTTAAACAGAAAGTATAGTTGCTCTGGGTCACTAGTT 4620
Qy 4621 CTGAATCTACAGATTCACAACTACAGAGGAACTTTTCCAAAAATTAAGGTGGCG 4680
Db 4621 CTGAATCTACAGATTCACAACTACAGAGGAACTTTTCCAAAAATTAAGGTGGCG 4680
Qy 4681 GAGTGTGATCTAGCAAGGATCAAACTTGTATTTCTTGTCAATTTCTGAAAA 4740
Db 4681 GAGTGTGATCTAGCAAGGATCAAACTTGTATTTCTTGTCAATTTCTGAAAA 4740
Qy 4741 CTACATATTAACAAGAACTTATATAGCAATTTGCAATTTCTCAATTAATTA 4800
Db 4741 CTACATATTAACAAGAACTTATATAGCAATTTGCAATTTCTCAATTAATTA 4800
Qy 4801 AAATGATTTAATCTATCTGGGAGAAAGTGCATAGATTAATTAACATTAATAGG 4860
Db 4801 AAATGATTTAATCTATCTGGGAGAAAGTGCATAGATTAATTAACATTAATAGG 4860
Qy 4861 AAATGAGCATCTGCAAGATTTTGTCTGTGCTGGGTTCTGAAAAATCCCTGTAAT 4920
Db 4861 AAATGAGCATCTGCAAGATTTTGTGTGTGCTGGGTTCTGAAAAATCCCTGTAAT 4920
Qy 4921 ACACAAAAATGACACTCTTGAGATCTGAACTAGAAAGCTCCAAAGATCATACAGAA 4980
Db 4921 ACACAAAAATGACACTCTTGAGATCTGAACTAGAAAGCTCCAAAGATCATACAGAA 4980
Qy 4981 TTCCAAAAATGCTGCTCCCAAGTCTAGAGGTTGCCCTCACTCTTGATCTCAAT 5040
Db 4981 TTCCAAAAATGCTGCTCCCAAGTCTAGAGGTTGCCCTCACTCTTGATCTCAAT 5040
Qy 5041 GGTTCACAGCAGATTAGCATTCAGATCTTAATGAAAAAGAGAGAGAGAGGCT 5100
Db 5041 GGTTCACAGCAGATTAGCATTCAGATCTTAATGAAAAAGAGAGAGAGAGGCT 5100
Qy 5101 TTGCTCTCTTAATTAATCCCATAGAGCAGAGACTTGTCTGTCACTTTGTGATCTTCC 5160
Db 5101 TTGCTCTCTTAATTAATCCCATAGAGCAGAGACTTGTCTGTCACTTTGTGATCTTCC 5160
Qy 5161 ACTTAACAGCACTGCTCATGAGATGTCATCCAGCATCAAGAACTGGGAGTGGGTC 5220
Db 5161 ACTTAACAGCACTGCTCATGAGATGTCATCCAGCATCAAGAACTGGGAGTGGGTC 5220
Qy 5221 CTGTGCTGCTGTATCTCAGAAAGGTTATGTGACCAAAAAGAAATCTTGGGGCA 5280
Db 5221 CTGTGCTGCTGTATCTCAGAAAGGTTATGTGACCAAAAAGAAATCTTGGGGCA 5280
Qy 5281 ACCAGAGCTCTTCAAGCCCTGAGCTGTCTGTATTTGTGTCACTCAAGATTTTCA 5340
Db 5281 ACCAGAGCTCTTCAAGCCCTGAGCTGTCTGTATTTGTGTCACTCAAGATTTTCA 5340
Qy 5341 GGAACCTCCAGAAATTAATTAATGAGGAGAGAGAGAACTGGAGTCCGTGAGGA 5400
Db 5341 GGAACCTCCAGAAATTAATTAATGAGGAGAGAGAGAACTGGAGTCCGTGAGGA 5400
Qy 5401 CTCGAGGAGCAGAGGCTGGTATTTGACCTGCTCTTCAATTTGTGAACAGAAAACTCT 5460
Db 5401 CTCGAGGAGCAGAGGCTGGTATTTGACCTGCTCTTCAATTTGTGAACAGAAAACTCT 5460
Qy 5461 TTAATTTCTAGGCTTACCTTATCTTATGATTAATGAGATTAATCAATACAGTCT 5520
Db 5461 TTAATTTCTAGGCTTACCTTATCTTATGATTAATGAGATTAATCAATACAGTCT 5520
Qy 5521 TTAAGAAATCATGATGAGATTTAAACAATGCTTAATGTTGTGATACCAAGTAAAA 5580
Db 5521 TTAAGAAATCATGATGAGATTTAAACAATGCTTAATGTTGTGATACCAAGTAAAA 5580
Qy 5581 AGACAGCATGATTAATCTCTTACGATCTTCCGTTTCTAGAGAGCAAAATATCA 5640
Db 5581 AGACAGCATGATTAATCTCTTACGATCTTCCGTTTCTAGAGAGCAAAATATCA 5640

Db 5581 AGACAGCATGATTAATCTCTTACGATCTTCCGTTTCTAGAGAGCAAAATATCA 5640
Qy 5641 TACCATCTATATGATGACAAACGTTTATGCTCTGATGTGATTAACAGAGTTCTC 5700
Db 5641 TACCATCTATATGATGACAAACGTTTATGCTCTGATGTGATTAACAGAGTTCTC 5700
Qy 5701 CTTTTCCTCCCAATTCCTTTTGGGCAATCAGAGCTGGGAGCTGCTCCCTAAG 5760
Db 5701 CTTTTCCTCCCAATTCCTTTTGGGCAATCAGAGCTGGGAGCTGCTCCCTAAG 5760
Qy 5761 AGCTCATGATGATGACACTCTGATGCTCTTATCTGAGATTTGAGTTAAATTTT 5820
Db 5761 AGCTCATGATGATGACACTCTGATGCTCTTATCTGAGATTTGAGTTAAATTTT 5820
Qy 5821 TTCTTCCACCTGGAGAGCTCTGCTCCATGTCATTTCTTGGATTCAGAGAACTGG 5880
Db 5821 TTCTTCCACCTGGAGAGCTCTGCTCCATGTCATTTCTTGGATTCAGAGAACTGG 5880
Qy 5881 CCTGATGAGCAAAAAGAACTCCGAAATTTGTTCTTAATGGCAGCACTCACTCTA 5940
Db 5881 CCTGATGAGCAAAAAGAACTCCGAAATTTGTTCTTAATGGCAGCACTCACTCTA 5940
Qy 5941 TTTTTCCTTAATTCATTTGCTTCTCATTTCTATCTGAGTTTGTAGTTAAATTTT 6000
Db 5941 TTTTTCCTTAATTCATTTGCTTCTCATTTCTATCTGAGTTTGTAGTTAAATTTT 6000
Qy 6001 TTTTTCAGCCCAATTTTGAATGATGATGATTTAACTTGAGAAATCACTCCCTACT 6060
Db 6001 TTTTTCAGCCCAATTTTGAATGATGATGATTTAACTTGAGAAATCACTCCCTACT 6060
Qy 6061 TTAACCCCTCTTAACATGATTAATGACATAGTGTGCTGGGTCCAAAGGCTGTGA 6120
Db 6061 TTAACCCCTCTTAACATGATTAATGACATAGTGTGCTGGGTCCAAAGGCTGTGA 6120
Qy 6121 AAAATGATCATGAGTACAGCTGCTGAGCTGCTCAATTCATTAATTAATTAACCC 6180
Db 6121 AAAATGATCATGAGTACAGCTGCTGAGCTGCTCAATTCATTAATTAATTAACCC 6180
Qy 6181 CCGGCAAAATTAATTCCTCTTTTAACTGATTAATTTCAATTTCTGCAATTAACAGAAATTA 6240
Db 6181 CCGGCAAAATTAATTCCTCTTTTAACTGATTAATTTCAATTTCTGCAATTAACAGAAATTA 6240
Qy 6241 TACTAATTTTAACTCTGAGCTCTTCAAGATGATTTAAAGAGCAATCTAATTAAC 6300
Db 6241 TACTAATTTTAACTCTGAGCTCTTCAAGATGATTTAAAGAGCAATCTAATTAAC 6300
Qy 6301 TCAATCAGCTGCTGATTTCTCCAAATTAAGCTTAATCTCAATCTCAGTTGAAATTT 6360
Db 6301 TCAATCAGCTGCTGATTTCTCCAAATTAAGCTTAATCTCAATCTCAGTTGAAATTT 6360
Qy 6361 GCATGAATATCTCTTTTAAACCAAGCCCTCACTCTCTATTTTCACTGAGACT 6420
Db 6361 GCATGAATATCTCTTTTAAACCAAGCCCTCACTCTCTATTTTCACTGAGACT 6420
Qy 6421 CCTCTCATCAATGTTTGCATCAACAAAGAACTTACCAAGATCTCCGAAAGAG 6480
Db 6421 CCTCTCATCAATGTTTGCATCAACAAAGAACTTACCAAGATCTCCGAAAGAG 6480
Qy 6481 AATGAATAGTTTACATTTGTATATCTCAGAGAACTTAATGATCCCAATACATAT 6540
Db 6481 AATGAATAGTTTACATTTGTATATCTCAGAGAACTTAATGATCCCAATACATAT 6540
Qy 6541 TCCCAACTTCAATTAATCTGCTGAGTGGCACTCAGGCTCAACCTCACTTACTTTTCT 6600
Db 6541 TCCCAACTTCAATTAATCTGCTGAGTGGCACTCAGGCTCAACCTCACTTACTTTTCT 6600
Qy 6601 CTGTTCTATGCTGAGCAATTCAGCTCAGACCAACCTTACCCAAACCTGTATACAA 6660
Db 6601 CTGTTCTATGCTGAGCAATTCAGCTCAGACCAACCTTACCCAAACCTGTATACAA 6660
Qy 6661 ATGCTTTCTAGGGGTTGCGAAAGCACTGAGTCTTATTTTAAAGGCAATCAGTGT 6720
Db 6661 ATGCTTTCTAGGGGTTGCGAAAGCACTGAGTCTTATTTTAAAGGCAATCAGTGT 6720

QY	6722	CAATTTCAGGTTTGGGACCTCATCAATCATCTCTCAACAAGATGAGCTGGCCA	6780
Db	6721	CAATTTCAGGTTTGGGACCTCATCAATCATCTCTCAACAAGATGAGCTGGCCA	6780
QY	6781	AATAGAAATTCGTAGTAATGAATTTTCTCTCATCTAATATAATGTGTGTTCTAATGCC	6840
Db	6781	AATAGAAATTCGTAGTAATGAATTTTCTCTCATCTAATATAATGTGTGTTCTAATGCC	6840
QY	6811	TACATTGTCCTTCAATTTTATATTTTCCATTTCATTCGAATCTACCATTTGCCATTAGGCTT	6900
Db	6811	TACATTGTCCTTCAATTTTATATTTTCCATTTCATTCGAATCTACCATTTGCCATTAGGCTT	6900
QY	6901	CTCATGATGATGATTCCTTCAATTTGATGATGAACTGTTATGAAAAGACATTGTGCTGTATG	6960
Db	6901	CTCATGATGATGATTCCTTCAATTTGATGATGAACTGTTATGAAAAGACATTGTGCTGTATG	6960
QY	6961	GAAATGGCACTTAGAGATATAAATGTGTAATGTGGTCTGTCTGCAATGACTGACACT	7020
Db	6961	GAAATGGCACTTAGAGATATAAATGTGTAATGTGGTCTGTCTGCAATGACTGACACT	7020
QY	7021	GAGTATTTTCTCACCACACAGGTCGGGCAATTTTCAACATCTCTAGGCAAGATCCCATTT	7080
Db	7021	GAGTATTTTCTCACCACACAGGTCGGGCAATTTTCAACATCTCTAGGCAAGATCCCATTT	7080
QY	7081	TCCTCTGGTTCATTAATGATGATATCTTTTTCCTGTCCAGATGACCAAGTCTGGTCATG	7140
Db	7081	TCCTCTGGTTCATTAATGATGATATCTTTTTCCTGTCCAGATGACCAAGTCTGGTCATG	7140
QY	7141	AGGGTGTCAACACCACTCTTTGTGTATCTGAATTCCTCCACCTGAGAGAAATTTCCAG	7200
Db	7141	AGGGTGTCAACACCACTCTTTGTGTATCTGAATTCCTCCACCTGAGAGAAATTTCCAG	7200
QY	7201	CCCAAGATAGAGTATTCATTCGGGTCACAGACCTGGCTAGATAGATGGGGGTGTTTGAT	7260
Db	7201	CCCAAGATAGAGTATTCATTCGGGTCACAGACCTGGCTAGATAGATGGGGGTGTTTGAT	7260
QY	7261	CCTAATGTTATTCCTCATATGACACAGAACTTGTGTGCACTGAGAGAGGTCAAGCTTC	7320
Db	7261	CCTAATGTTATTCCTCATATGACACAGAACTTGTGTGCACTGAGAGAGGTCAAGCTTC	7320
QY	7321	AGAGTCACAAAGAACCTGATTTCAAACTGGATTTGAGAGACCCCACTTTTGATAGGTGA	7380
Db	7321	AGAGTCACAAAGAACCTGATTTCAAACTGGATTTGAGAGACCCCACTTTTGATAGGTGA	7380
QY	7381	CTTATTTCTGCGAGTCTCTGATCTCTCTCTTTAATAGAGACAGTAATCCCACTAG	7440
Db	7381	CTTATTTCTGCGAGTCTCTGATCTCTCTCTTTAATAGAGACAGTAATCCCACTAG	7440
QY	7441	CAGGTCGTGGGAGAAATCAGAGATCAAAACAGCTGTGTATCAATCTGGTTCGTGTTCC	7500
Db	7441	CAGGTCGTGGGAGAAATCAGAGATCAAAACAGCTGTGTATCAATCTGGTTCGTGTTCC	7500
QY	7501	AGGGTCATCAGACTGGGGTTTCTGAGCATGATTCAACTCCAGTCTTGGGTACAGAA	7560
Db	7501	AGGGTCATCAGACTGGGGTTTCTGAGCATGATTCAACTCCAGTCTTGGGTACAGAA	7560
QY	7561	CTGACACCAATTCACGACGCTGAGAGACCTCCTTGCTACAGCAGACCTGAGCTTCAAG	7620
Db	7561	CTGACACCAATTCACGACGCTGAGAGACCTCCTTGCTACAGCAGACCTGAGCTTCAAG	7620
QY	7621	GGGCTGACGTGATGTTTCCCTTGTGCGCGCTGACAGAAACGGCGTTGTGCTCTGGGCTC	7680
Db	7621	GGGCTGACGTGATGTTTCCCTTGTGCGCGCTGACAGAAACGGCGTTGTGCTCTGGGCTC	7680
QY	7681	CTGGGCTGCCGATCGCAGAGACGCTGTCTCCATTCATCATCTTCAACTGGTGCAGGCC	7740
Db	7681	CTGGGCTGCCGATCGCAGAGACGCTGTCTCCATTCATCATCTTCAACTGGTGCAGGCC	7740
QY	7741	GACTTCCTCTTCTTAAAGGCGCACTTATATGTTCCCGTTAGCGCTCATTAATCTCGC	7800
Db	7741	GACTTCCTCTTCTTAAAGGCGCACTTATATGTTCCCGTTAGCGCTCATTAATCTCGC	7800

QY	7801	CATCCCACTCCAAAATCTCTAGCTCGGTGATGACCTTCCCTACTTTATAGGGCTAAG	7860
Db	7801	CATCCCACTCCAAAATCTCTAGCTCGGTGATGACCTTCCCTACTTTATAGGGCTAAG	7860
QY	7861	ATGCTGAGCGCCATCAGCACCGAGCGCTGCCTGTCCATCTCTGAGCCCATCTGAGACAC	7920
Db	7861	ATGCTGAGCGCCATCAGCACCGAGCGCTGCCTGTCCATCTCTGAGCCCATCTGAGACAC	7920
QY	7921	TGCCCGCCCCCAATATACCTGTGATTCGGGTATATGTGTCTCTGTCTGTGGGCCCTGTCCCTG	7980
Db	7921	TGCCCGCCCCCAATATACCTGTGATTCGGGTATATGTGTCTCTGTCTGTGGGCCCTGTCCCTG	7980
QY	7981	CTGCGAGATATCTGAGTGTGATTTCTGTACTTCCGTATTAGTGTGATCTGTGTT	8040
Db	7981	CTGCGAGATATCTGAGTGTGATTTCTGTACTTCCGTATTAGTGTGATCTGTGTT	8040
QY	8041	TGCTGTGAAAGCTCAGATTTCAATTAACAATCCCGGTGGCTTTTATATGTGTGTTCTC	8100
Db	8041	TGCTGTGAAAGCTCAGATTTCAATTAACAATCCCGGTGGCTTTTATATGTGTGTTCTC	8100
QY	8101	TGTGGGTCACAGCTGTGATCTGTGCTGATGAGATTCCTGTGTGATTCGCCGAGATGCGCGTG	8160
Db	8101	TGTGGGTCACAGCTGTGATCTGTGCTGATGAGATTCCTGTGTGATTCGCCGAGATGCGCGTG	8160
QY	8161	ACCAAGGCTGTACGTGACCCATCTCTCCATCAGATGTGTGTTCTTCTCTGTGTGGCTGCC	8220
Db	8161	ACCAAGGCTGTACGTGACCCATCTCTCCATCAGATGTGTGTTCTTCTCTGTGTGGCTGCC	8220
QY	8221	TTTGGCATTCAGTGGGCCCTGTGTTTCCAGATTCACCTGGAATTGAAAAGTCTTATTTTGT	8280
Db	8221	TTTGGCATTCAGTGGGCCCTGTGTTTCCAGATTCACCTGGAATTGAAAAGTCTTATTTTGT	8280
QY	8281	CATGTCATCTAGTTTCCATTTTCTGTGCGGCTCTTAAACAGCAGATGTGCCAATCCCATCATTT	8340
Db	8281	CATGTCATCTAGTTTCCATTTTCTGTGCGGCTCTTAAACAGCAGATGTGCCAATCCCATCATTT	8340
QY	8341	TACTTCTTCTGTGGCTCTCTTTAGGACGTCMAAATAGCAGAACTGTAAAGCTGTTCTC	8400
Db	8341	TACTTCTTCTGTGGCTCTCTTTAGGACGTCMAAATAGCAGAACTGTAAAGCTGTTCTC	8400
QY	8401	CAGAGGGCTCTGACAGACACGCCCTGAGGTGATGAAAGTGTGAGGGTGTCTTCTCAGAGAA	8460
Db	8401	CAGAGGGCTCTGACAGACACGCCCTGAGGTGATGAAAGTGTGAGGGTGTCTTCTCAGAGAA	8460
QY	8461	ACCTGTGAGCTGTGGGAGACGATTGTGAGCAGTGAAGAAACCTGTGCCCTGTCAAGC	8520
Db	8461	ACCTGTGAGCTGTGGGAGACGATTGTGAGCAGTGAAGAAACCTGTGCCCTGTCAAGC	8520
QY	8521	AGGACTTTGAGAGCAATGCTGCCCCCTGACACCTTGAACAATATATGACTTTTCTTAGCC	8580
Db	8521	AGGACTTTGAGAGCAATGCTGCCCCCTGACACCTTGAACAATATATGACTTTTCTTAGCC	8580
QY	8581	TTCTGCGCTCAGAAATGTCTCAGGGTCTCCCAAGGCCCTTACCA 8622	
Db	8581	TTCTGCGCTCAGAAATGTCTCAGGGTCTCCCAAGGCCCTTACCA 8622	

```

RESULT 2
US-10-741-601-5719
; Sequence 5719, Application US/107441601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michèle et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CE001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5719
; LENGTH: 396287
; TYPE: DNA

```


Qy	1993	1992
Db	333973 AAGATCTTGGCATAACATTTTATATGCAAGTACAAATTGGCAACC	3340323
Qy	1993 -----CAATCCATTAAATATGTATTAATATATATTTATATACATGG	2037
Db	334033 AGCCCAATGCCCATCAATCAATTAAGTGTATTAAGAAATGTGGTTATGTACCATGG	3340923
Qy	2038 AATACAACAGCCGCAATTAATAAAGATAAATAATGATGACATTCACAGCAATCTAATGGAAT	2097
Db	334093 AATACTTACTCACCACATTAAGAGAAAGAAATATATGCAATTCACAGCACTGTGATGAAT	334153
Qy	2098 TEGAGACCCCTTAATTCCTAAGTGGGGTACTCAGGAATGGAACCAACATCATATGTTCT	2157
Db	334153 TGGAGACTATTAATCTTAAGTGAAGTAATCAGGAATGGAACCAACATCGATGTTCT	3342123
Qy	2158 CACTTACAAGTGGGGGCTAAGCTGTGAGAGACAGAGGCTATGATATATTAATGAACTC	2217
Db	334213 CACTTAATATGTGGAGCTAAGCTATGTGGTACACAAAGGCAT -GAGGTATCGTTGACTT	3342717
Qy	2218 TEGGCACTTGAGGGGAGAGATGGAAGAGAGGCGAGGAGATTAAGACTTACCAATGGGTAC	2277
Db	334272 TGGAGACTCAATGTAAGAAAGGTG ---GGGTGGCAAGAGGTTAAAGACTACATTTGGGTAC	3343282
Qy	2278 AGTGTACACTGCTCAGGTATATGGGTGCAACCAAAATCTCAGAAATTCACATTAAGAACTT	2337
Db	334329 AGTGTACACTGCTCAGGTATATGGGTGTTACCAAACTCAGAAATTCACCACTTAAGAAATTT	334388
Qy	2338 ATCTCATGGAAGCAACACACCACTGTTCCCAAAATCCCAATGAATAATTAATATAA	2397
Db	334389 TTTATATGTAAACCAACACACCACTGTTCCCAAAACCTTATTAATACACACAAABAAAT	334444
Qy	2398 T 2398	
Db	334449 T 334449	

```

RESULT 3
US-10-450-826-47/c
? Sequence 47, Application US/10450826
? Publication No. US20040101818A1
? GENERAL INFORMATION:
?   APPLICANT: Ji, Darren
?   APPLICANT: Axelrod, Douglas W.
?   APPLICANT: Cook, Jonathan S.
?   APPLICANT: Jaiswal, Neelam
?   APPLICANT: Eistein, Richard
?   APPLICANT: Houghton, Adam
?   APPLICANT: Mertz, Lawrence
?   TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
?   FILE REFERENCE: 044921-5039-NO
?   CURRENT APPLICATION NUMBER: US/10/450, 826
?   CURRENT FILING DATE: 2003-06-18
?   PRIOR APPLICATION NUMBER: US 60/255,882
?   PRIOR FILING DATE: 2000-12-18
?   PRIOR APPLICATION NUMBER: US 60/285,691
?   PRIOR FILING DATE: 2001-04-24
?   NUMBER OF SEQ ID NOS: 149
?   SOFTWARE: PatentIn Ver. 2.1
?   SEQ ID NO 47
?   LENGTH: 66494
?   TYPE: DNA
?   ORGANISM: Homo sapiens
?   FEATURE:
?   OTHER INFORMATION: Genbank Accession No. AC006957
US-10-450-826-47

```

Query Match	16.2%	Score 1401;	DB 17;	Length 66494;
Best Local Similarity	77.7%	Pred. No. 0;		
Matches 2034;	Conservative	0;	Mismatches 375;	Indels 208;
			Gaps	20;
09	15	GTCACTTTAATCCAAAACCGAAGATAT-----ACAAAAAGAAACTATTAGCC	68	

Db	64729	GTCAACCTTAATTCACAAACCGAGGAAGACATACACAAACAAAAGAAAACCTATAGACC	64670
Qy	69	AGTACCACTGATGATATATACATGACAGAAATCCCAACAAATATCTAGTAACTCAATCCA	128
Db	64669	AATATCCCTAATGAAATATAGATGCACAAAATCCACAGAAAATATTGTGAACCAATCCA	64610
Qy	129	ACAGCATATCAAGAAATATCCACCATTTGTCAAGTGGTTCTATACCGAGGGTG-CAGG	187
Db	64609	ACAAATATATGAAAGAAATATCCACATGATCAAGTGGGTTCTATACCAAGATGCCA	64550
Qy	188	ATAGGTTAACTACACAGTCAATTAATGATGATCATCATTAACAGAAATTAAAAACA	247
Db	64549	ATAGTTTAACTATGTATAGTCAATTAATGTGATACCCATTAACAGATGAAAAACA	64490
Qy	248	AAATCAGATGATCATCTCAATAGATCTGAAAAAGCATTTGACAAATCTAACATTTCTT	307
Db	64489	AAATCAGATGATCATCTAAATTAATGATTCAGAAAAAGCATTTGACAAACCCAGATCCCTT	64430
Qy	308	TATGATTTAAACCTTCAGCAAAATCCATATGAAAGACATACCTTAATGTATTAAGC	367
Db	64429	TATGATTTAAACCTTTGGCAAAATTCAGCATTAAGGAGACATATCTTAAGGTATTAAGC	64370
Qy	368	CATATATGAGGAGCCACAGCAAACTTTATCTGAATGGGAGAAAGTTGAAAAACCTTGC	427
Db	64369	CATCTATGACAAACCCACAGCCAACTATATATGAAT-GGAGAAAGTTTAAAGCATTTCTC	64311
Qy	428	CCTGAGAACTGSAACAGACAAAGATG-CTACTTTCACCACTTCTTATTCACATAGTACT	486
Db	64310	TATTAAGAACTGSAACAGAAAGATGCCACTTTACACACTTCTTATTCACATAGTACT	64251
Qy	487	GGAGTTTAAAGCAGACCAATCAACAAGAAAGAAATCAAGGCAACCCAAATCAATTA	546
Db	64250	GGAGTTTAAAGCAGACCAATCAACAAGAAAGAAATCAAGGCAACCCAAATCAATTA	64191
Qy	547	AGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATATGATTGATATCTAGAAAACTTAA	606
Db	64190	AGAGAAAGTCAAACTGCTCACTGTCACAGAAATGATGATTGTATCTAGAAAACTTAA	64131
Qy	607	AGACTCATCAAGAAAGCTCCTTAAGCTGATACATTAATTCAGTAAAGTTTCAGATACAA	666
Db	64130	AGCTCATCCAAAAGCTCCTTAATCTGATGATGAATTAATCAATTAAGTTTCAAGATGTA	64071
Qy	667	ACTAATATGTCACAAATCAGTAGCACTGCTATACACCAACAGTGACCAAGCTGAGATCA	726
Db	64070	AATCAATGTATACAAATCAGTAGCTCTGCCCTACACCAACAGAAATCAAGTTGAGATCA	64011
Qy	727	AATCAAGACTCAAAACCTTTTACATAGCTGTAAA-----AAAT	767
Db	64010	AATCAATCACTCAACCCCTTTTCAATAGCTGCAAAAAATTAATTAATTAATTAATTAAT	63951
Qy	768	ACTTAAGATATTTCTTACCCAGAGAGGTGAAGACCTCTTACAGAGAAATCTCAAAACAC	827
Db	63950	TTCTTGGAAATATCTTACCAAGAGGTGAAGACCTCTTACAGAGAAATCTCAAAACAC	63891
Qy	828	AGCT-----GACATCATAGATGACACAAACAGTGAACACATCCCATGCTCATGGATGG	883
Db	63890	TGCTGAAGAAATCAACGAGGACCAAAATTAATTAAGAAACACATCCCATGCTCATGGATGG	63831
Qy	884	GTAGATCATATTTGTGAAAAATGACATATTTGCCAAAGCAATCTTCAAGTTCAATGCA	943
Db	63830	GTAGATCATATTTGTGAAAAATGACATATTTCTTCAAGAACTTACAAATTCATATGCA	63771
Qy	944	TTCCACCAAAATTCATCATCTCTTCAAGAACTTAAAGAAAAACAATTTCTAAATTC	1003
Db	63770	TGCCCAACAACTTCACTCAATTAATTTCTTCAAGAACTTAAAGAAAAACAATTTCCCAAGTC	63711
Qy	1004	AT-----ATGAAACAAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT	1050
Db	63710	ATGGAATTTTAAAGAACATGGCTAAATATCTGGAACCAAAAAAAGAGCCCAATATGTCAAA	63651
Qy	1051	GCAAGACTTAAAGAAACAAATCTGAGGCACTACATTCACCATTTTCAACTATATAC	1110

Db 63650 GCAAG-CTAAGCAAAAAGAAACAATCTGGAG---ACACATGCTGACTTCAAACTATAC 63595
 Qy 1111 TACAAGGCTATATACCAAAAATCATGCACTGACATTAATCTAGGACATAGACCA 1170
 Db 63594 CA-TAGGCAATGTCATCAAAAATCATGCTGATTAATTAAC-AGGCACTAGACCA 63537
 Qy 1171 TGGAAAAGAGAGAAATCCAGAAATTAAGCCAAATTAATTAATAGCAATGATTTTGC 1230
 Db 63536 TGGAAACAGATAGAAACCCAAATAAAGCCAAACCTTACAGCCAACTGATCTTCAAC 63477
 Qy 1231 AAAGCAACAAAACATTAAGTGGGGAAGAACATCTTGAATTACAAATGCTGAGAT 1290
 Db 63476 AAAGCAACAAAACATTAAGTGGGGAAGAACCACTTATTAACAATGCTGAGAT 63417
 Qy 1291 TATTGGCAAGCCACATGTTGGAAGAAATGAATCCCTGCTCTCACTTAATCAAA 1350
 Db 63416 AATTGGCAAGCTACATGTAAGAAATTAATTAATCTGGAT-CCCTCATCTCTCACCGTATTA 63358
 Qy 1351 AATTGATCAAGATGATCAAGAACTTAATCTGAGACCTTAACATTAATTAATCTAGA 1410
 Db 63357 AATCAACTCAGATGATCAAGAACTTAATCTGAGACCTTAACATTAATTAATCTAGA 63298
 Qy 1411 AGATTAATCAAGAAATGCTTCTAGACATTCATTAAGCAAGACTTCAATGCGCAAGAA 1470
 Db 63297 AGATTAATCAAGAAATGCTTCTAGACATTCATTAAGCAAGACTTCAATGCGCAAGAA 63238
 Qy 1471 CCCAAAGTAAATGCAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1530
 Db 63237 CCCAAAGTAAATGCAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 63178
 Qy 1531 CTTTGGCAGCAAAAACATCATTTAGCAGAGCAACAGCAACCCAGCTGAGAGAA 1590
 Db 63177 C-TGCGACAGCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 63127
 Qy 1591 AATCTTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1650
 Db 63126 ---TTTGCACAACTTTCATCTGACAAAGAACTTAATTAATTAATTAATTAATTA 63071
 Qy 1651 ACAATTCAGAAAGAAAGCAAAATCCCATGAAGTGGGCTTGAAGCACTGAATG 1710
 Db 63070 ACAATTCAGAAAGAAAGCAAAATCCCATGAAGTGGGCTTGAAGCACTGAATG 63016
 Qy 1711 ACAATTCAGAAAGAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1768
 Db 63015 ACAATTCAGAAAGAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 62956
 Qy 1769 ACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1828
 Db 62955 ACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62896
 Qy 1829 AATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1888
 Db 62895 AATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62836
 Qy 1889 AACACTTTTACACTGCTGCTGGAATGTAATTAATTAATTAATTAATTAATTAATTA 1948
 Db 62835 GACACTTTTACACTGCTGCTGGAATGTAATTAATTAATTAATTAATTAATTAATTA 62776
 Qy 1949 GAAATTTCTTAAGAACTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2000
 Db 62775 GAGATTTCTTAAGAACTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62716
 Qy 2001 ----- 2000
 Db 62715 TATCTAACCAAGCAAGAAAGTCAATTAACAAAAGATTAATTAATTAATTAATTA 62656
 Qy 2001 ----- 2000
 Db 62655 AATGATATTTTAT 62596
 Qy 2001 ---TTAAT 2057
 Db 62595 TAT 62536

Qy 2058 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2117
 Db 62535 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62476
 Qy 2118 GGGGTAACTCAGAAATGAAACCAAAATCATATATGTTCTCACTTAAGAGTGGGCTTA 2177
 Db 62475 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62416
 Qy 2178 GCTGT---GAGACACGAAGGAT-AGATGATTAATTAATTAATTAATTAATTAATTAATTA 2233
 Db 62415 GCTGTGAGAGAGACGCAAGGATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 62356
 Qy 2234 AGGATGAAGAGAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2293
 Db 62355 AGGATGAAGAGAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62296
 Qy 2294 GTGATGGGTGACCAAAATCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2353
 Db 62295 GTGATGGGTGACCAAAATCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62236
 Qy 2354 ACCACTGTTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2413
 Db 62235 ACCACTGTTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62177
 Qy 2414 ACAGATTTAAAAAAGTTCATGTTTCAAGTTTATTA 2450
 Db 62176 ATTGCATGAGTATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 62140

RESULT 4
 US-10-087-192-418/c
 ; Sequence 418, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; PRIOR FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 418
 ; LENGTH: 209484
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(209484)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-418

Query Match 16.2%; Score 1395; DB 13; Length 209484;
 Best Local Similarity 78.1%; Pred. No. 0;
 Matches 1989; Conservative 0; Mismatches 385; Indels 173; Gaps 19;

Qy 1 TGTATGAAGCCATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 58
 Db 93988 TGTATGAAGCCATGTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 93929
 Qy 59 ACTATAGACCAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118
 Db 93928 ACTATAGACCAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 93869
 Qy 119 ACCCAATTCACAGCATATC--AAGAAGTAATTCACCATGTCAGAGTGGTTTCATTAAC 176
 Db 93868 ACCCAATTCACAGCATATC--AAGAAGTAATTCACCATGTCAGAGTGGTTTCATTAAC 93809

QY 177 AGGGGTGAGATGTTGTTAATACATACAGAGTCAATTAATGTGATCATCATTAACAGA 236
 DB 93808 AGGATATGAGATGTTAATACATCATTAAGTCAATTAATGTGATCAACCATGAATAGA 93749
 QY 237 ATTAAAAAACAATATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAAATC 296
 DB 93748 ATTAAAAAACAATATCATGATCATCTCAATAGATGATTAATAAGCATTTGACAAAATC 93689
 QY 297 TAACATTTCTTTATGATTTAAAACTTCAGCAAAAATCGACATAGAAAAGCATACCTTAAT 356
 DB 93688 CAGATCCCTTTATGATTTAAAACTTCGAAAAGCAGCATAGAAAAGCATCTTTAG 93629
 QY 357 GTATTAATAAGCCATATATGACGAGCCACAGCAAACTTATATCTGAATGGGAAAAAGTTG 416
 DB 93628 GTATTAATAAGCCATATATGACAAATCCAGCCAACTTATATCTGAATGGGAAAAAGTTG 93569
 QY 417 AAAACATTTGCTCGAGAACTGGAACAAGAGATG-CTACTTTGACCACTTCTATTC 475
 DB 93568 AAAGCATTTCCCTTGGGAACTGGAACAAGAGATGCTCACTTTGACCACTTTATTC 93509
 QY 476 AACATAGTAGGAAGTTTACGAGCAATCAGACAAAGAAAGAAATCAAGGGCAC 535
 DB 93508 AACAAAGTAGGAAGTTTACGAGCAATCAGACAAAGAAAGAAATTAAGGGCAC 93449
 QY 536 CAATCATTAAGGAAGATCAAACTGCTCTGTTCACTGATGATGATTTACTTA 595
 DB 93448 CAATCATTAAGGAAGATCAAACTGCTCTGTTGATGATGATGATTTACTTA 93389
 QY 596 GAAACCTTAAGATCTAGCAAGAGTCTCTAGAACTGATACATAATTCATTAAGT 655
 DB 93388 GAAACCTTAAGATCTAGCAAGAGTCTCTCAAGAGTCTGTAATTTGGCAAGT 93329
 QY 656 TCAAGATACAACTTAATGTACACAAATCAGTAGCACTGATACCAACAGTGAACA 715
 DB 93328 TCAAGATACAAATTAATGTACAAATCAGTAGCACTGATACCAACAGTGAACA 93269
 QY 716 GCTGAGATCAATCAAGAACTCAACCTTTCAATTAAGTCTGTA- 762
 DB 93268 GCTGAGATCAATTAAGAACTCAACCTTTCAATTAAGTCTGTA- 93209
 QY 763 AAAATCTTAAGATTTCTTACCAAGAGTGAAGAGCTCTACAAAGAACTCA 822
 DB 93208 AAATTACTTAAGATTTCTTACCAAGAGTGAAGAGCTTCTGCAAGGAACTCA 93149
 QY 823 AACACAGCTGACAT- - -CATAGATGACAAACAAGTGAAGACATCCCATGCTATG 878
 DB 93148 AACCTGCTGAATTAACCATGATGACAAACAAGTGAAGACATCCCATGCTATG 93089
 QY 879 GATGGGTAAGATCAATTTGTGAAAAATGACCAATTTGCCAAAGCAATCTACAGTTCAA 938
 DB 93088 GAAGGGTAAGATCAATTTGTGAAAAATGACCAATTTGCCAAAGCAATCTCA 93029
 QY 939 TGCATTTCCCAACCAAAATATCATCATTTCTTACAGAACTAGAAAAAAACAATTTCAA 998
 DB 93028 TGCATTTCCCAATTAACCAACCAATTTCTTACAGAACTAG-AAAAAAACAATTTCAA 92970
 QY 999 AATTCTATGGAACAAACAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 1058
 DB 92969 AATTCTATGGAAC- - - - -CAAAAGAGCCGCTGCTAGCAAAACAAGCT 92924
 QY 1059 TAGCAAAAAGAACAAATCTGAGGACATCACTTACCAATCTTCAACTATCTACAAAGC 1118
 DB 92923 AAGCAAAAAGAACAAATCTGAGGACATCACTTACCAAGTTCAAAAGATCTTAAAGGC 92864
 QY 1119 TATATCAACCAAAACATCATGCACTGACATTAAGCACTAGCACTAGCACTAGCACTAG 1178
 DB 92863 CATATCAACCAAAAGCATGCTAGTATTAATAAGCACTAGCACTAGCACTAGCACTAG 92804
 QY 1179 AAGAGAAATCCAGAAATTAAGCAAAATTAATATAGCACTAGTTTTTGAACAAAGCAA 1238
 DB 92803 AATAGAAATCCAGAAATTAAGCAAAATTAATATAGCACTAGCTTTGCAAAAGCAA 92744

QY 1239 CAAAAATTAAGTGGGAAAAAGCATTTCTAGTTAAACAATGTGCTGACATTAATGGCA 1298
 DB 92743 CAAAAATTAAGGAGGAAAAAGCACTTTATTTAAACAAAGGTGCTGGGATTAATGGCA 92684
 QY 1299 AGCCATGAGGAAGATGAAATCGATCCCTTGTCTCTCACTTAATPACAAAATTTGATA 1358
 DB 92683 AGCCATGAGGAAGATGAAATTTGAT-CTTGCTCTCTCACTTAATPACAAAATTTGATA 92625
 QY 1359 CAAGATGATCAAAAGACTTAATCTGAGACCTTAACCATTAATAATTTCTAGAAATTA 1418
 DB 92624 CAAGATGATCAAGACTTAATACTAGACCTTAATCATGATCATGATGATGATGATGAT 92565
 QY 1419 TCAGAAAAATGCTTCTGACATTTCACTTGAAGCAAAAGCTTATGAGCCAAAGAACCA 1478
 DB 92564 TTGAAAAACCTTCTGACATTTGACATTTGCAAAAGCTTATGACAAAGAACCAAG 92505
 QY 1479 TAAATGCAACAAAAACAATAATTAATGATGAGACTTAATTAATTAATAAAAGCTTTG 1538
 DB 92504 CAATGCAACAAAAACAATAATTAATGATGAGACTTAATTAATAAAAGCTTTGCA 92445
 QY 1539 CAGCAAAAACATCATTTAGCAGACAAACAGCAACCAACCAAGTGAAGAAAACTTTCA 1598
 DB 92444 CAGTAAAGAAATTAACAGCAGAG-TTACAAAGCAACCAACAGTGAAG-AAAACTTTCA 92387
 QY 1599 CAATTAAGCATCTGATTAAGATTAATTTCCGAAATTCACAGAACTCAACAAATCA 1658
 DB 92386 CAATTAAGCATCTGATTAAGATTAATTTCCGAAATTCACAAAGTCTCAACAAATCA 92327
 QY 1659 GCAAGAAAGAAAGCAAAATCCATGAAAGAGTGGGCTTAAGGACATGAATGACAAATCT 1718
 DB 92326 GCAAGAAAGAA- - - - -AAATTCCTCAAAAGATGAGGCTTAAGGACATGAATGACAAATCT 92271
 QY 1719 CAAAGAAAGATTAACAATG- - - - -CCAAACAACAGAAAAAATGCTTAACATCACTAATGA 1776
 DB 92270 CAAAGAAAGATTAACAATGAGTCCAGACATATGAAAAATGCTCAACATCACTAATGA 92211
 QY 1777 TTAGGAAATGTAATC- - - - -AACATGTAATGCAATACCACTTCTCTGCAAGATGCT 1834
 DB 92210 TCAGGAAATGCAAAATCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 92151
 QY 1835 CATATTTAAAAATCTAAAAATTAATGATGTTGGTGGGCTGCTGATTAAGAACT 1894
 DB 92150 CATATTTAAAAAT- - - - -AAATTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 92092
 QY 1895 TTACACTGCTGCTGAGATGTAATCTTGGCAACAATAAGAAAAAGTGGAAAT 1954
 DB 92091 TTACACTGATAGTGGAAATGTAATCTATGATCACTCTGAAAAAGTGGAAAT 92032
 QY 1955 TCTTAAGAACTAAAGTA- - - - - 1973
 DB 92031 CTTTAAAGAACTAAAGATCTCTTTGATCTAGCAATTTCTAGTATTTAATCTAG 91972
 QY 1974 - - - - - 1973
 DB 91971 AGAAAAAGATCATTTATCAAAAAAAGATACCTGATCATGATCTTTGTAGCAGACAAAT 91912
 QY 1974 - - - - -GATGACCAATTTGATCCAGCAATCCATTAATATGATTA- - - - - 2015
 DB 91911 TCGCAATTTGAAAAATATGGAACAGCCCAATCCATCAATCAACAGTGAATTAAT 91852
 QY 2016 TATATATTTATTTATACATGAAATCACTGAGCCATTAATAAAGATTAATAATGATGCA 2075
 DB 91851 GTGATATTTATTTATTAACATGAAATCACTGAGCCATTAATAAAGATTAATAATGATGCA 91792
 QY 2076 TTACAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2135
 DB 91791 TTGAGCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91732
 QY 2136 AAAACCAACATCATATTTCTCACTTAACAAGTGGGCTTAAGCTGTAAGCAACAAG 2195
 DB 91731 AAAACCAACATCTATGTTCTCATTTATGATGAGGCTTAATTAAGAGGCGCTAAG 91672
 QY 2196 CAT-AGATGATTAATGAATCTGAGGAACTTGAAGGGAAGATGAGAGAGCGAGGG 2254

```

Db 91671 CATAGATTAATACATTGACCTTGGGAGCTTGGGGGAAAGGT-GAGGGGTGGCGAGGA 91613
Qy 2255 ATAAAAAGCTACACAAATGGGTAACAGTATACATGCTCAGGTGATG--GGTGCACCAAA 2311
Db 91612 ATAAAAAGCTACACAAATGGGTAACAGTATGCTGCTCAGGTGATGATGACCAAA 91553
Qy 2312 TCTCGAAATTAACCTTAAGAACTATCATGGAAGCAACGCACTGTTCCCAAA 2371
Db 91552 TCTCGAAATTAACCTTAAGAAATTTATTCATGTAACCAATGCCAGCTGTTCCCAAA 91493
Qy 2372 TCCCAATGAAATTAATTAATTAAT 2398
Db 91492 ATTATTTGAAATTAATTAATTAAT 91466

RESULT 5
US-10-331-053-4/c
; Sequence 4, Application US/10311053
; Publication No. US20040197778A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001100
; CURRENT FILING DATE: 2002-12-26
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 209484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(209484)
; OTHER INFORMATION: n = A,T,C or G
US-10-331-053-4

Query Match 16.2%; Score 1395; DB 18; Length 209484;
Best Local Similarity 78.1%; Pred. No. 0;
Matches 1989; Conservative 0; Mismatches 385; Indels 173; Gaps 19;

Qy 1 TGTATGAAGCAATGTCACCTTATATCAAAACGAGAAAGATAT--CAAAAAGAA 58
Db 93988 TCTATGACGCCAGCATGCGCTTATACCAAAACGAGGAGATACACAAAAGGAA 93929
Qy 59 ACTATAGACCAAGTACCACTGATGAATATACATGACAGAAATCCCAACAAATATAGCTA 118
Db 93928 ACTACAGACCAATATCCCTGATGAACATAGATCAAAAATCTTCAACAAATATAGTGA 93869
Qy 119 ACCCAATCCACAGCATATC--AAGAAGATATCCACCATTTGTCAGAGTGGTTCTATAC 176
Db 93868 ACCCAATCCACAGCATATC--AAGAAGATATCCACCATTTGTCAGAGTGGTTCTATAC 93809
Qy 177 AGGGGTGAGATAGTATACATACCAAGTCAATTAATGTATACATCATATAACAGA 236
Db 93808 AGGGATATAGATGTTTAACTATACATTAAGTCAATTAATGTATACCACTGATATAGA 93749
Qy 237 ATTAATAACCAAAATACATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATC 236
Db 93748 ATTAATAACCAAAATACATGATCATCTCAATAGATGATATAAAGCATTTGACAAATC 93689
Qy 297 TAAACATTTCTTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGACATACCTTAAT 356
Db 93688 CAGTATCCCTTTATGATTAATAACCTTCAGCAAAAGCAGCATAGAAAGGACATCTTAAG 93629
Qy 357 GTAATTAAGGCAATATATAGACGACCAAGCAAACTTATATGATGGGAAAGTTG 416
Db 93628 GTAATTAAGGCAATATATAGACCAAGCAGCAATCTTATACGATGGGAAAGTTG 93569
Qy 417 AAAACATTTCCCTGGAACCTGGAACAAGACAGATG--CTACTTTCACCACTTCTATTC 475

```

```

Db 93568 AAAGCATCCCTCGGAACGTGAACAAGATGTCACCTTTCACCACTTTATATTC 93509
Qy 476 AACATAGTGTGAAAGTTTATACGAGCAATACAAAGAAATCAAGGGCAC 535
Db 93508 AACAAAGTACTGAAGTCTTATACGAGCAATACAAAGAAATCAAGGGCACATC 93449
Qy 536 CAATCAATTAAGAGAGAGTCAAACTGTCCTGTTCACTGATATATGATTTATACCTA 595
Db 93448 CAATCAATTAAGAGAGAGTCAAACTGTCCTGTTCACTGATATATGATTTATACCTA 93389
Qy 596 GAAACCTTAAGACTATCCAGAAAGTCTTGAATCTGATATCAATTAATTCAGTAAAT 655
Db 93388 GGAACCTTATGATGATCCAAAAAGCTCCAGAACTGTTAAATGAAATTTGGCAAGTT 93329
Qy 656 TCAGATTAACAATTAATGATACAAATGATAGGACCTGCTATACCAAGAGTACCA 715
Db 93328 TCAGATTAACAATTAATGATAGCAATGATAGGACCTGCTATACCAAGAGTACCA 93269
Qy 716 GCTGAATCAATCAAGAACTCAAACTTTTACAAATGATGCTGTA-- 762
Db 93268 GCTGAATCAATTAAGAACTCAAACTTTTACAAATGATGCTGTA-- 93209
Qy 763 AAAATCTTAAGATATTTCTTACCAAGAGTGAAGACCTTACAAAGAAATCTCA 822
Db 93208 AAATCTTAAGATATTTCTTACCAAGAGTGAAGACCTTACCAAGAGAACTCA 93149
Qy 823 AACAGCTGACAT----CATGATGACCAAAAGTGAAGACATCCCATGCTCATG 878
Db 93148 AACAGCTGAAATTAACATGATGACCAAAAGTGAAGACATCCCATGCTCATG 93089
Qy 879 GATGGTGAATCAATATTTGTAATAATGACATATTTGCCAAAGAACTCAATCA 938
Db 93088 GAAGGTGAATCAATATTTGTAATAATGACATATTTGCCAAAGAACTCAATCA 93029
Qy 939 TGCATTTCCACCAAAATTCATCATCTTCTTACAGAACTGAAAGAAATCAATTTCA 998
Db 93028 TGCATTTCCACCAAAATTCATCATCTTCTTACAGAACTGAAAGAAATCAATTTCA 92970
Qy 999 AATTCATTAAGAAACCAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1058
Db 92969 AATTCATTAAGAAACCAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 92924
Qy 1059 TAGCAAAAGAAACCAATCTGAGGATCATCAATTCATTTCAATTAATTAATTAATTAAT 1118
Db 92923 TAGCAAAAGAAACCAATCTGAGGATCATCAATTCATTTCAATTAATTAATTAATTAAT 92864
Qy 1119 TATATACCAAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1178
Db 92863 CATATGACCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 92804
Qy 1179 AAGAGAAATCCAAATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1238
Db 92803 AAGAGAAATCCAAATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 92744
Qy 1239 CAAAAACATTAAGTGGGAAAGAGACATTTCTAGTTAACTAATGATGCTGATTAATTTGCA 1298
Db 92743 CAAAAACATTAAGGAGAAAGAGACATTTCTAGTTAACTAATGATGCTGATTAATTTGCA 92684
Qy 1299 AGCCATGTTGAAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
Db 92683 AGCCATGTTGAAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 92625
Qy 1359 CAAGATGATCAAAAGCTTAATCTGAGACCTTAATTAATTAATTAATTAATTAATTAATTA 1418
Db 92624 CAAGATGATCAAAAGCTTAATCTGAGACCTTAATTAATTAATTAATTAATTAATTAATTA 92565
Qy 1419 TCAGAAATATGCTTACGATTTCACTTATGAGCAAGATTTCAATGAGCAAGAACTTTCG 1478
Db 92564 TTGGGAAACCTTTACGATTTCACTTATGAGCAAGATTTCAATGAGCAAGAACTTTCG 92505
Qy 1479 TAAATGCAACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1538
Db 92504 CAATGACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 92445

```


NUMBER OF SEQ ID NOS: 84
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1691139
TYPE: DNA
ORGANISM: Human
US-10-067-514-1

Query Match 16.0%; Score 1375.6; DB 14; Length 1691139;
Best Local Similarity 77.1%; Pred. No. 6,1e-313;
Matches 1993; Conservative 0; Mismatches 409; Indels 182; Gaps 19;

QY 1 TGTATGAAGCCAAAGTCACTTTATACCAAAACGAGAAAGATAT--CAAAAAAGAA 58
DB 1325552 TCTATGAAGCCAGCATCCCTAGTACCAAAACGAGAAAGATATCAAAAAAGAA 1325611
QY 59 ACTATGACACGATACCACTGATGATATACATGCAAGAAATCCCAACAAATATCTAGCTA 118
DB 1325612 ACTACAGATCAATTCCTGATTAACATAGTGTAAATCTTACCAAAATCTAGCTG 1325671
QY 119 ACCCAATCCACACGATATCAAGAAATATCCCAATTTGCAAGTGGTTTCATACCGAG 178
DB 1325672 ACTCAATCCACACATATCAAAAGATATCCCAATGATCAAGTGGTTTCATACCGAG 1325731
QY 179 GGGTGCAG-GATAGGTTAACTACACAAGTCAATTAATGTATCATCTACATTAACAGAA 237
DB 1325732 GGGTGCAGAGTGGTTTAATGTATACACAAGTCAATTAATGTATCACCACTAAACAGAA 1325791
QY 238 TTAATAAACAATAATCACATGATCATCTCAATAGATGCTGAATAAAGCAATTTGACAAATCT 297
DB 1325792 TTAATAAACAATAATTCATGATCATCTCAATAGTGAAGAAAGCAATTTGACAAATCTC 1325851
QY 298 AACATTTCTTTATGATTTAAACCTTCAGCAAAATCGACATAGAAAGACATATCCTTAATG 357
DB 1325852 AGCATCCCTTTATGATTTAAAGCTCTCAGCAAAATCACATATCAAGGACATATCTTAATG 1325911
QY 358 TAAATAAAGCATATATGAGCGAGCCACAGCAACATTAATCTGAATGGGAAAGTTGA 417
DB 1325912 TAAATAAAGCATATCTATGACCAACCCACAGCAACGTAATCTGAATGGGAAAGTTGA 1325971
QY 418 AACCATTTGCTCTGAGATGAGACAGACAGACAGAGATG-CTATCTTCAACACTTCTATTC 476
DB 1325972 AAGATTTCCCTCTGAGATGAGACAGACAGACAGATATGCTCTCAGCACTCTTCTTCA 1326031
QY 477 ACATAGTAGTGAAGTTTAAAGCCAGACATCAGACAGAGAAAGAAATCAAGGCAACC 536
DB 1326032 ACATAGTAATGGAAGTCTAGCAAGACCAATCAGACAGAGAGGAATAAAGGCAATCC 1326091
QY 537 AAATCAATTAAGGAAGTCAAACTGTCCTGTTCACTGATGATGATGTATTAACCTAG 596
DB 1326092 AAATCGGTAAAGGAAGTCAAACTGTCATGTTGCTGATGATGATGTATTAACCTAG 1326151
QY 597 AAAACCTTAAGACTCATCCGAGAAAGCTCTGAGATGATCATTAATTTCTAGTAAAGTTT 656
DB 1326152 AAAACCTTAAGACTCTCTCAGCAAGCTCTTGAATCTGAATTAAGTAAAGTTT 1326211
QY 657 CAGATATCAAACTTAATGATACAAATCAAGTGAAGTCTATACCAACAGTGAACAG 716
DB 1326212 CTGGAATCAAGATTAATGATACAAATCAAGTGAAGTCTTCTATATACCAACAGTGAACAG 1326271
QY 717 CTGAGATCAAAATCAAACTCAAACTTTTAAATAGCTGTAA-----AAA 765
DB 1326272 GGGAGAAATCAAACTCAAACTCAAACTTTTAAATAGCTGTAA-----AAA 1326331
QY 766 ATACTTAAGATATTTCTTACCAAGAGTGAAGGACCTCTACAGAGAAATCTACAAAC 825
DB 1326332 ATACTTAAGATATTTCTTACCAAGAGTGAAGGACCTCTACAGAGAAATCTACAAAC 1326391
QY 826 ACAGCTGA-----CATCATAGTGAACAAACAGTGAAGAACATCCCATGCTCATGAT 881
DB 1326392 ACTGCTTGAAGAGATCATAGCAATTAACAAATGAGGAACATCCCGTGAATCATGAT 1326451
QY 882 GGGTAGAATCAATTTGTGAATAATGACCATATTTGCCAAAGCAATCTACAGTTCAATGC 941

DB 1326452 GGGTAGAATCAATTTGTGAATAATGACCATATTTGCCAAAGCAATCTACAGTTCAATGC 1326511
QY 942 AATTCACCAAAAATATCATCATCTTCTTACAGAACTAGAAAAAACAATTTCAAAAT 1001
DB 1326512 AATTCACCAAAAATATCAATCATCTTCTTACAGAACTAG--AAAAACAATTTCAAAAC 1326570
QY 1002 TCATATGGAACACACCAAAAAA-----CCCCGATAGCCCAAGCAAGCTTAG 1061
DB 1326572 TCATATGGAAC-----CAAAAAAGCTGATAGCCCAAGCAAGCAAG 1326616
QY 1062 CAAAAAGCAAAATCTGAGGCAATCATATTAACCATCTTCAAACTATATCTCAAGGCTAT 1121
DB 1326617 CAAAAAGTCAAAATCTGAGGCAATCATCTTCAAGCTATTTCAAACTATATCTCAAGGCTAT 1326676
QY 1122 AATCACCAAAAATCATGAGCACTGACATTAATACTAGGCACTAGAACCAATGAGAAAGAG 1181
DB 1326677 AGTCAACCAAAAATGAGCAATGATCTGTATTAATAATAGCATATATAGCAATGAGAAAGAT 1326736
QY 1182 AGAGATTCAGAAATTAAGCCAAATTAATATAGCCAACTGATTTTGAACAAACAA 1241
DB 1326737 AGAGAACCCGAGATTAACCCAAATATCTTACAGCCAACTGATCTTCAACAAAGCAACAA 1326796
QY 1242 AAAATATTAAGTGGGAAAA--GACATCTAGTTAACAAATGCTGAGATTAATGGCAAG 1300
DB 1326797 AAACATTAAGTGGGAAAAAGATTAACCTTTCAACAAATGCTGAGATTAATGGCTAG 1326856
QY 1301 CCACATGTGAAGAAATGAACCTGATCCCTGTCTCTCACTTAATACAAAAATGATATCA 1360
DB 1326857 CCACATATGAGAGATGAATGATATGATCC--TATCTCTCACCGTATACAAAAATCACTCA 1326914
QY 1361 AGATGATCAAGACTTAATCTGAGACTTAACCAATTAATAATTTAGAGATTAATC 1420
DB 1326914 AGATGATTAAGGCTTAATCTTAAGAGCTGAACATAT--GAATTTTGAAGATTAATCTT 1326973
QY 1421 AGAAAAATGCTTCTAGACATTCACCTTAGGCAAGACTTCAATGGCAAGAACCCAAAGTA 1480
DB 1326973 AGATGATTAAGGCTTAACTTAAGAGCTGAACATAT--GAATTTTGAAGATTAATCTT 1326973
QY 1481 AATGCAACAAACAAATAATTAATAGATAGACTTAATTAATCA--TAAAGGCTTTTGGC 1539
DB 1327034 AATGCAATTAATAAACAAGATTAATAGCTGAGACCTCATTAATCTTACAGCTTTTGCAG 1327093
QY 1540 AGCAAAAACATCAATTAGCAGAGCAACAGACAAACCAACGAGTGAAG--AAAATCTTCA 1598
DB 1327094 GGCAAAAGCAAGCTGACAGAGTAACAGACAAACCAACGAGTGAAGAAAAATCTTCA 1327153
QY 1599 CAATCTAAGCATCTGACTAAGACTAATATCCGGAATCCCAAGGAATCAAAACAAATCA 1658
DB 1327154 CAATCTAATCTCTGACAAAGG--GTAGTATCCAGAAATCTACAGAGACCCCAACAAATCA 1327212
QY 1659 GCAAGAAAGAAAGCAACATCCCATGAAGAGTGGGCTAAGGACATGAATAGCAATCTCT 1718
DB 1327213 GTAGAAAGAAAGCAACATCCCATGAAGAGTGGGCTAAGGACATGAATAGGCAATTTCA 1327272
QY 1719 CAAAAGAAATATACAAATGAGCCAAAC--AGAAAAAATGCTTAACATCACTAATGA 1776
DB 1327273 CAAAAGAAATATATCAATATGAGCAAGCAACATATGAATAAATGCTCAACATCAATATGA 1327332
QY 1777 TTAGGGAATATGTAATGAACA--CTGTAATGCAATTCACCTTATCTCTCAAGAAATGCT 1834
DB 1327333 TCAAGGAATATGCAATTAACAAACCAAAATGTAATACACCGTATCTTCAAGAAATGCT 1327392
QY 1835 CATTAATTAATAATCT--AAAAATTAATAGATGTTGAGTGTCTGTGTATTAAGAACCA 1892
DB 1327393 CATTAATTAATAATTTTAAATAACAGTATGTTGGATGGAAGGGGTGATCAGGAAACA 1327452
QY 1893 CTTTACACTGCTGTGTGGGAATGTAACTTGCGCAACCATATGAGAAACAGTGTGAAA 1952
DB 1327453 CTTTACACTGCTGTGTGGGAATGTAACTATGATCAGCACTTAATGGAAGAAAGTGTGGA 1327512
QY 1953 TTTCTTAAGGACATTAAGATGATGCAATTTGATCAGCAATCCCAT----- 2001

Db	1327513	TTCCTTAAAGAACATAAAGTAGAACTACACAATTGATCCAGACGTGCCACTCTAGTATC	1327572
Oy	2002	-----	2001
Db	1327573	TACCAGAGAAAAGAGTCATTATTGGAAAAAGACCTTGTAACGTATGTTATAGCA	1327632
Oy	2002	-----	2001
Db	1327633	GCACATTTCACAATTGCCAAAACCTGTGGAACTTACCCAAATGTCATCGTCAATTAAGTGG	1327692
Oy	2002	-TAAATATGTATAAATATATATATATTATATACATGAATACAACTCACCATTA AAAAG	2060
Db	1327693	ATTAAGAAGAACTGTGTGTACAGATATATATACAAATGAATACTACACAGCATGAAAAAG	1327752
Oy	2061	AATTAATAATGATGACATTGACACGCATCTTGAATGGAAATGGAGACCCCTTATTTCTAATGGGG	2120
Db	1327753	AATGATATTACACGACATTTGACAGTGCCTGGAGAGATGGAGACATATTATTTCTAAGTGA	132812
Oy	2121	GTAATCTCAGAAATGAAAAACCAAATCATATGTTCTCATCTTACAAAGTGGGGGCTAAGCT	2180
Db	1327813	GTAATCTCAGAAATGAAAAACCAAATCATATGTTCTCATCTGATATGTGGGATCTTAAGCT	1327872
Oy	2181	GTGAGACACGAAAGCAT-AGAAATATTAATGAACTCTGGGACCTTGAAGGGAAAGATG	2239
Db	1327873	ATGAGGACAAAGATATTAAGAAATGATCCAATGACCTTTGGGGACTTGGGGGAAAGAGTGG	1327932
Oy	2240	GAAAGAGAGCGAGGATTAAGAACTACACAAATGGGTAAAGGTACACATGCTCAGGTGATG	2299
Db	1327933	GAGGGGGGTGTGGGATTAAGAAATTAAGAAATGTGTGACGTGTATACGTCTTGGGTGATG	1327992
Oy	2300	GGTGACACAAATCTCAGAAATTAACCATTAAGAACTTATCCATGGAAGCAA--ACACCA	2357
Db	1327993	GGTGACACAAATCTCAGAAATTAACCATTAAGAACTTATTAATCTAAATATACCA	1328052
Oy	2358	CCTGTTCCCCAAAATCCCAATGAATTA AAAATTAATTAATTAATGAATTTAATTTCAACG	2417
Db	1328053	CCTGTACCTTATTAATCTTAATGAAAAAAATTTTATTA AAAAGTAATTAATGAATTTAAGTCANA	1328112
Oy	2418	AATT 2421	
Db	1328113	GTTT 1328116	

```

RESULT 8
US-10-419-723-1
; Sequence 1, Application US/10419723
; Publication No. US20040014099A1
; GENERAL INFORMATION:
; APPLICANT: Greterdottir, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reynisdottir, Sigríður Th.
; APPLICANT: Thorleifsson, Gudmar
; APPLICANT: Gulcher, Jeffrey
; TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
; TITLE OF INVENTION: METHODS OF TREATMENT
; RILE REFERENCE: 2345, 2010-005
; CURRENT APPLICATION NUMBER: US/10/0419, 723
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 10/255, 120
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 10/067, 514
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 09/811, 352
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
; TYPE: DNA
; ORGANISM: Homo sapiens
Query Match 15.0%; Score 1375.6; DB 16; Length 1691139

```

Best Local Similarity 77.1%; Pred. No. 6.1e-313;
Matches 1993; Conservative 0; Mismatches 409; Indels 182; Gaps 19;

QY	1	IGTATGAAGCCAAATGTCATCTTTAATATACCAAAACAGAGAAAGATAT- -CAAAAAAATA	58
Db	1325552	TCATGAAGCCAGCATCACCTTACTGACCAAAACAGAGAAAGATATACCAAAAAAATA	1325611
QY	59	ACTATGACGACGATCCACTGATGAAATATACATGACAGAAATCCCAACAAATCTAGCTA	118
Db	1325612	ACTACGATCAATATCTTGGATTAACATAGATGGTAAATCTTTACAAATATCTAGCTG	1325671
QY	119	AACCAATCCAAACAGATATCAAGAAAGATATCCACCATTCGATGGGTTCTTACACG	178
Db	1325672	ACTGATATCAACACATATCAAAAAGATATCCACCATGATCAAGTGGGTTCTTACACG	1325731
QY	179	GGGTGTCAG- GATAGGTTAAACATACACAAAGTCAATAAATGTTATATCATCATATAACGAA	237
Db	1325732	GGGTGACAGATGTTTAAATGTTACCAAGTCAATAAATGTTATATCATCATATAACGAA	1325791
QY	238	TTAAAAACAAAATCAGATGATCATCTCTCAATAGATGCTGAAAAAGCATTTGACAAAATCT	297
Db	1325792	TTAAAAACAAAATTCATGATTCATCTCAATAGATGCGAAAAAGCATTTCAACAAAATCC	1325851
QY	298	AAACATTTCTTATGATTTAAACCTTCAGCAAAATCGACATGAAAGACATACCTTATG	357
Db	1325852	AGCATCCCTTATGATTTAAACCTTCAGCAAAATCGATACAAAGGACATACATTAATG	1325911
QY	358	TAAATAAAGCATATATGACGAGCCACAGCAAAACATTTATCTGATATGGGAAAAGTTGA	417
Db	1325912	TAAATAAACATATCTATGACCAACCCACAGCCAAAGTAAATGCGAAAAGTTGA	1325971
QY	418	AAACATGTCCTCTGAGAACTGGAACAGACAGAGATG- TATCTTTCACCACTTCTATTTCA	476
Db	1325972	AAGAATTCCTCTGAGAACTGGAACAGACAGATGTCCTCTCACACTCTCTTCTTCA	1326031
QY	477	ACATGATGTCGAAGTTTATAGCCAGCAATACACAAAGAAATCAAGGGACCC	536
Db	1326032	ACATGATGTCGAAGTCTTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	1326091
QY	537	AAATCAATAAAGAGAGATCAAACTGTCCTCTGTTCACTGATGATATGTTGATATACCTAG	596
Db	1326092	AAATGCTAAAGAGAGATCAAACTGTCCTCTGTTGCTGATGATATGATTTTACCTTG	1326151
QY	597	AAAACTTAAAGATCATCCAGAAAGTCTCTAGAACCTGATATCATTAATTCATGTAAGTTT	656
Db	1326152	AAAACTTAAAGATCTCTCCAGCAAGCTCTAGAACCTGATTAAGAAATTCAGAAAGTTT	1326211
QY	657	CAGATATCAAACTAAATGTCACAAATAGTAGCATGCTTATATATCAACAGTACCAAG	716
Db	1326212	CTGATATCAAACTAAATGTCACAAATAGTAGCTTATATATCAACAGTACCAAG	1326271
QY	717	CTGAGAAATCAATCAAGAACTCAAAACCTTTTCAATAGCTGTAAA- ------AAA	765
Db	1326272	GGGAAATCAATCAAGAACTCAAAACCTTTTCAATAGCTGTAAAATTAATTAATA	1326331
QY	766	ATACTTAAGATATTTCTTATCCCAAGAGTAGAAGACTCTTACAGAGAAAATCTACAAAC	825
Db	1326332	ATACTTAAGATATTTCTTATCCCAAGAGTAGAAGACTCTTACAGAGAAAATCTACAAAC	1326391
QY	826	ACAAGTGA- ----CATCATAGATGACAAACAAGTGAACAACAATCCCATGCTATGAT	881
Db	1326392	ACTGCTGAAGAGAAATCATAGCAATATCAAAACAAATGGGAACAATCCCGATATGAT	1326451
QY	882	GGGTGATATCAATTTTGTGAATATGACATATTTGCCAAAAGCAATCTTACAGATTCATATGC	941
Db	1326452	GGGTGATATCAATTTTGTGAATATGACATATTTGCCAAAAGCAATCTTACAGATTCATATGC	1326511
QY	942	AATTCACCACAAAATATCATCATATTTCTTACAGAACTAGAAAAAAACAAATCTTAAAT	1001
Db	1326512	AATCCCATCAAAATATCATCATATTTCTTACAGAAATTTAG- -AAAAACAATTTCTAAAC	1326570
QY	1002	TCATATGACACACCAAAAAAAAAAAAAAAAACCCGCATATACCAAGCAAGACTTATG	1061

QY 7196 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGT 7255
 Db 1 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGT 60
 QY 7256 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCACTAGAGAGGTGAG 7315
 Db 61 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCACTAGAGAGGTGAG 120
 QY 7316 GCTTCAGATCAACAAGAACTGATTTTCAACTGAGATTGAGAGACCCCACTTTGATA 7375
 Db 121 GCTTCAGATCAACAAGAACTGATTTTCAACTGAGATTGAGAGACCCCACTTTGATA 180
 QY 7376 GGTGACTTATCTCTGAGTCTCTGATCTCTCTCTTTAATGAGACAGTAATCCCA 7435
 Db 181 GGTGACTTATCTCTGAGTCTCTGATCTCTCTCTTTAATGAGAGAGTAATCCCA 239
 QY 7436 CATGCGAGGGTGTGGGAGAAATCAGAGATCAACAGCTGGTATCACTGTGGTCTG 7495
 Db 240 CATGCGAGGGTGTGGGAGAAATCAGAGATCAACAGCTGGTATCACTGTGGTCTG 239
 QY 7496 TTTCCAGGGTCAACAGCTGGGTTCTGAGCATGATTTCAACATCCCACTTGGGTA 7555
 Db 300 TTTCCAGGGTCAACAGCTGGGTTCTGAGCATGATTTCAACATCCCACTTGGGTA 359
 QY 7556 CAGAACTGACCAATCAACGAGCTGAGAGAGACTCTTCTCAAGCAGACCTGAGCT 7615
 Db 360 CAGAACTGACCAATCAACGAGCTGAGAGAGACTCTTCTCAAGCAGACCTGAGCT 419
 QY 7616 TCACGGGGCTGACGTGATCTGTTTCCCTGTGCGCTGACAGAAACGGGTGTGCT 7675
 Db 420 TCACGGGGCTGACGTGATCTGTTTCCCTGTGCGCTGACAGAAACGGGTGTGCT 479
 QY 7676 GGTCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCACTCACTCACTCACTGCTG 7735
 Db 480 GGTCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCACTCACTCACTCACTGCTG 539
 QY 7736 CGGCGACTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7795
 Db 540 CGGCGACTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 599
 QY 7796 TCCGCGATCT 7855
 Db 600 TCCGCGATCT 659
 QY 7856 TTAAGATGCTGAGCGGATCAACGAGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCT 7915
 Db 660 TTAAGATGCTGAGCGGATCAACGAGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCT 719
 QY 7916 ACCACTGCGCGCGCGCGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7975
 Db 720 ACCACTGCGCGCGCGCGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 779
 QY 7976 CCCTGCTGCGAGATCT 8035
 Db 780 CCCTGCTGCGAGATCT 839
 QY 8036 CTGTTTGTGTAAGATGATGATTTTCAATTAACATGCGGTGCTGTTTATGTTGAG 8095
 Db 840 CTGTTTGTGTAAGATGATGATTTTCAATTAACATGCGGTGCTGTTTATGTTGAG 899
 QY 8096 TTCTCTGTGGTTCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8155
 Db 900 TTCTCTGTGGTTCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 959
 QY 8156 CGCTGACAGGCTGTAAGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8215
 Db 960 CGCTGACAGGCTGTAAGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
 QY 8216 TGCCCTTTGGCAATTCAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8275
 Db 1020 TGCCCTTTGGCAATTCAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1079
 QY 8276 TTTGTATGTCATCTAGTTTCAATTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 8335

Db 1080 TTTGTATGTCATCTAGTTTCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1139
 QY 8336 TCATTTACTTCT 8395
 Db 1140 TCATTTACTTCT 1199
 QY 8396 TTTCTCAGAGGGCTCTGACAGACAGCTGAGGTGATGAAGTGAAGGTGGTCTCTCT 8455
 Db 1200 TTTCTCAGAGGGCTCTGACAGACAGCTGAGGTGATGAAGTGAAGGTGGTCTCTCT 1259
 QY 8456 AGGAAACCTTGAAGCTGTGCGGAGAGAGATGAGAGATGAGAGAACTTCTGCTCT 8515
 Db 1260 AGGAAACCTTGAAGCTGTGCGGAGAGAGATGAGAGATGAGAGAACTTCTGCTCT 1319
 QY 8516 CAGACAGACTTTGAGAGCAATGCTGCTGCACTCTGCAATTAATATGATTTTCT 8575
 Db 1320 CAGACAGACTTTGAGAGCAATGCTGCTGCACTCTGCAATTAATATGATTTTCT 1379
 QY 8576 TAGCTTCTGCTCAGAAATG 8596
 Db 1380 TAGCTTCTGCTCAGAAATG 1400

RESULT 10

US-10-225-567A-673
 ; Sequence 673, Application US/10225567A
 ; Publication No. US20030113798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lifespan Biosciences
 ; APPLICANT: Brown, Joseph P.
 ; APPLICANT: Burner, Glenna C.
 ; APPLICANT: Roubh, Christine L.
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 1920-4-4
 ; CURRENT APPLICATION NUMBER: US/10/225,567A
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/257,144
 ; NUMBER OF SEQ ID NOS: 2292
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 673
 ; LENGTH: 1400
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-225-567A-673

Query Match

15.9%; Score 1374.6; DB 15; Length 1400;

Best Local Similarity 99.3%; Pred. No. 0; Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 7196 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGT 7255
 Db 1 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGT 60
 QY 7256 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCACTAGAGAGGTGAG 7315
 Db 61 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCACTAGAGAGGTGAG 120
 QY 7316 GCTTCAGATCAACAAGAACTGATTTTCAACTGAGATTGAGAGACCCCACTTTGATA 7375
 Db 121 GCTTCAGATCAACAAGAACTGATTTTCAACTGAGATTGAGAGACCCCACTTTGATA 180
 QY 7376 GGTGACTTATCTCTGAGTCTCTGATCTCTCTCTTTAATGAGACAGTAATCCCA 7435
 Db 181 GGTGACTTATCTCTGAGTCTCTGATCTCTCTCTTTAATGAGAGAGTAATCCCA 239
 QY 7436 CATGCGAGGGTGTGGGAGAAATCAGAGATCAACAGCTGGTATCACTGTGGTCTG 7495
 Db 240 CATGCGAGGGTGTGGGAGAAATCAGAGATCAACAGCTGGTATCACTGTGGTCTG 299
 QY 7496 TTTCCAGGGTCAACAGCTGGGTTCTGAGCATGATTTCAACATCCCACTTGGGTA 7555

Db 300 TTTCAGGGGTACCAAGACTG666GTTTCTGAGCATGAGATTCAACCATCCAGTCTTG6GTA 359
 QY 7556 CAGAACTGACACCAATCAACGAGCGTGAAGAGACTCTTGCTACAGACAGACCTTGAGCT 7615
 Db 360 CAGAACTGACACCAATCAACGAGCGTGAAGAGACTCTTGCTACAGACAGACCTTGAGCT 419
 QY 7616 TCACAGGGGGCTGACGTGATGTTTCCCTTGCTGCGCTGACAGGAAAGCGGTGTGCTCT 7675
 Db 420 TCACAGGGGGCTGACGTGATGTTTCCCTTGCTGCGCTGACAGGAAAGCGGTGTGCTCT 479
 QY 7676 GGCCTCTGAGGCTGCGAGTCGAGAGACGCTGCTCATATGATCTCAACCTGAGTGC 7735
 Db 480 GGCCTCTGAGGCTGCGAGTCGAGAGACGCTGCTCATATGATCTCAACCTGAGTGC 539
 QY 7736 CGGCGGACTTCTCTTCTTCTTACGCGGCAATTAATATGTTGCGCGCTTACGCTCATATA 7795
 Db 540 CGGCGGACTTCTCTTCTTACGCGGCAATTAATATGTTGCGCGCTTACGCTCATATA 599
 QY 7796 TCCGCGCATCCCATCTTCCAAAATCTCTGAGTCGTGATGAGTCTTCCCTACTTTATAGGCC 7855
 Db 600 TCCGCGCATCCCATCTTCCAAAATCTCTGAGTCGTGATGAGTCTTCCCTACTTTATAGGCC 659
 QY 7856 TTAACATGCTGAGGCGCATGACGACGAGGCGTGCATGCTGCTGAGGCCCATCTGCT 7915
 Db 660 TTAACATGCTGAGGCGCATGACGACGAGGCGTGCATGCTGCTGAGGCCCATCTGCT 719
 QY 7916 ACCACTGCGCGCGCGCCAGATACCTGATCATGCTGATGCTGCTGCTGCGCGCGCTG 7975
 Db 720 ACCACTGCGCGCGCGCCAGATACCTGATCATGCTGATGCTGCTGCTGCGCGCGCTG 779
 QY 7976 CCGTCTGCGGAGATATCTGAGATGAGATGTTCTGATCTTCTGTTATAGTGTCTGAT 8035
 Db 780 CCGTCTGCGGAGATATCTGAGATGAGATGTTCTGATCTTCTGTTATAGTGTCTGAT 839
 QY 8036 CTGTTTGTGTGAACGTCAGATTCTTACATCGCGTGTGCTGTTTATAGTGTG 8095
 Db 840 CTGTTTGTGTGAACGTCAGATTCTTACATCGCGTGTGCTGTTTATAGTGTG 899
 QY 8096 TTCTCTGTGGGTCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8155
 Db 900 TTCTCTGTGGGTCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
 QY 8156 CGCTGACAGGCTGTAGTGAACATCTCTCTCAAGTGTGTGCTCTCTCTCTGCTGCTG 8215
 Db 960 CGCTGACAGGCTGTAGTGAACATCTCTCTCAAGTGTGTGCTCTCTCTCTGCTGCTG 1019
 QY 8216 TGCCCTTTGGGATCAGTGGGCGCTGTTTCCAGATCCACTGGAATTGGAAGTCTTAT 8275
 Db 1020 TGCCCTTTGGGATCAGTGGGCGCTGTTTCCAGATCCACTGGAATTGGAAGTCTTAT 1079
 QY 8276 TTGTGTCATGTGATCTAGTTCATTTTCTGCTGCTGCTTAAAGAGTGCACACCCA 8335
 Db 1080 TTGTGTCATGTGATCTAGTTCATTTTCTGCTGCTGCTTAAAGAGTGCACACCCA 1139
 QY 8336 TCATTTACTTCTTGTGTGGGCTCTTTAGGAGCGCTCAAAATAGGCAAGAACTGAAAGCTG 8395
 Db 1140 TCATTTACTTCTTGTGTGGGCTCTTTAGGAGCGCTCAAAATAGGCAAGAACTGAAAGCTG 1199
 QY 8396 TTCTCCAGAGGCTCTGAGAGGACGCTGAGTGAATGAAGGTGAAGGAGGCTTCTC 8455
 Db 1200 TTCTCCAGAGGCTCTGAGAGGACGCTGAGTGAATGAAGGTGAAGGAGGCTTCTC 1259
 QY 8456 AGGAAACCTGAGCTGTGCGGAAGCAGATTGAGAGAGTGAAGAACTCTGCGCTGT 8515
 Db 1260 AGGAAACCTGAGCTGTGCGGAAGCAGATTGAGAGAGTGAAGAACTCTGCGCTGT 1319
 QY 8516 CAGACAGAGCTTTAGAGCAATGCTGCGCTGCGACCTTGAATTAATGATGATTTTCT 8575
 Db 1320 CAGACAGAGCTTTAGAGCAATGCTGCGCTGCGACCTTGAATTAATGATGATTTTCT 1379
 QY 8576 TAGGCTTCTGCTCAGAAATG 8596
 Db 1380 TAGGCTTCTGCTCAGAAATG 1400

RESULT 11
 US-10-322-696-4/c
 ; Sequence 4, Application US/10322696
 ; Publication No. US20040166490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Malandro, Marc
 ; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
 ; CURRENT APPLICATION NUMBER: US/10/322,696
 ; FILING DATE: 2003-10-17
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 100944
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-322-696-4

Query Match 15.9%; Score 1370; DB 17; Length 100944;
 Best Local Similarity 78.1%; Pred. No. 2,2e-312;
 Matches 2013; Conservative 0; Mismatches 355; Indels 210; Gaps 22;

QY 1 TGTATGAAGCCAAATGTCACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 58
 Db 46567 TGTATGAAGCCAAATGTCACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 46508
 QY 59 ACTATGAAGCCAAATGTCACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 118
 Db 46507 ACTATGAAGCCAAATGTCACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 46448
 QY 119 ACCAATCCAAAGCATATCAAGAGATTAATCCCAATTTGCAAGTGGTTTCATACAG 178
 Db 46447 ACCAATCCAAAGCATATCAAGAGATTAATCCCAATTTGCAAGTGGTTTCATACAG 46388
 QY 179 GGGTGAGG--ATAGTTAATACCAAGTCAATTAATGATATCATCATTAACAGAA 237
 Db 46387 GGGTGAGG--ATAGTTAATACCAAGTCAATTAATGATATCATCATTAACAGAA 46328
 QY 238 TTAATAAACAATAATCAATATCATCTGATGCTGTAAGAAAGATTTGACAAATCT 297
 Db 46327 TTAATAAACAATAATCAATATCATCTGATGCTGTAAGAAAGATTTGACAAATCT 46269
 QY 298 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 357
 Db 46268 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 46210
 QY 358 TAATTAAGCCATATATGAGCGAACCAAGCAAAATTAATATCTGAATGGGAAAGTTGA 417
 Db 46209 TAATTAAGCCATATATGAGCGAACCAAGCAAAATTAATATCTGAATGGGAAAGTTGA 46150
 QY 418 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 477
 Db 46149 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 46092
 QY 478 CATATGATGGAAGTTTATGAGCGAACCAAGCAAAATTAATATCTGAATGGGAAAGTTGA 537
 Db 46091 CATATGATGGAAGTTTATGAGCGAACCAAGCAAAATTAATATCTGAATGGGAAAGTTGA 46032
 QY 538 AATCAATTAAGAGAGAACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
 Db 46031 AATCAATTAAGAGAGAACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 45975
 QY 598 AAACCTTAAGAGCTATCCGAAAGCTCTGAGAACTGATATCAATTAATCTGTAAGTTTC 657
 Db 45974 AAACCTTAAGAGCTATCCGAAAGCTCTGAGAACTGATATCAATTAATCTGTAAGTTTC 45915
 QY 658 AGATACAAACTAAATGATACAAATGATGATGATGATGATGATGATGATGATGATGATGAT 717
 Db 45914 AGATACAAACTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 45865


```

QY 718 TGAGATCAAAATCAAGAACTCAAAACCTTTAACAATAGCTGTA--AAAAATCTTAAGA 775
DB 45864 TGAGATCAAAATCAAGAACTCAAAACCTTTAACAATAGCTTAACAATTAATTAAGGA 45805
QY 776 ATATTCTTAACCAAGAGGTGAAGACCTCTCAAGAGAAAATCAAAAACACAGCT---- 831
DB 45804 ATATACCAACCAAGAGGTGAAGACCTCTTAAGAGAAAATTAATTAACAGTGTGAAA 45745
QY 832 GACATCATATGATGACAAACCAAGTGAAGACATCTCCATGCTCATGATGGGTAGAAATC 891
DB 45744 GACATCATATGATGACAAACCAAGTGAAGACATCTCCATGCTCATGATGGGTAGAAATC 45685
QY 892 AATATTGTGAATGACATATTTGCAAGAACTCAAGAACTCAAGTTCATGCAATTTCCACC 951
DB 45684 GATCTGTGAATTTGACCACTGCAAGAACTCAAGTTCATGCAATTTCCATC 45625
QY 952 AAAATATCATCATCTTTCTTCAAGAACTCAAGAACTCAAGTTCATGCAATTTCCACC 1011
DB 45624 AAAATATCATCATCTTTCTTCAAGAACTCAAGAACTCAAGTTCATGCAATTTCCACC 45566
QY 1012 CAACCAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 1071
DB 45566 C-----CAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 45520
QY 1072 AAATCTGAGAGCATCATTAACCATCTTCAAACTAATCTAAGGCTATTAATCACCAAA 1131
DB 45519 AAATCTGAGAGCATCATTAACCATCTTCAAACTAATCTAAGGCTATTAATCACCAAA 45460
QY 1132 ACATCATGAGCATCATTAACCATCTTCAAACTAATCTAAGGCTATTAATCACCAAA 1191
DB 45459 ACATCATGAGCATCATTAACCATCTTCAAACTAATCTAAGGCTATTAATCACCAAA 45400
QY 1192 GAAATTAAGCCCAATTAATTAATGAGCACTGATTTTGAAGCAAAACCAAAACCAATTAAG 1251
DB 45399 GAAATTAAGCCCAATTAATTAATGAGCACTGATTTTGAAGCAAAACCAATTAAG 45345
QY 1252 TGAGGAGAGAGCATCTTAATTAACAATGAGTGTGATTAATGAGGCAAGCATGTGA 1311
DB 45344 TGAGGAGAGAGCATCTTAATTAACAATGAGTGTGATTAATGAGGCAAGCATGTGA 45285
QY 1312 AGAATGAACCTGATCCCTTGTCTCTCACTTAATTAACAATGAGTGTGA 1371
DB 45284 AGAATGAACCTGATCCCTTGTCTCTCACTTAATTAACAATGAGTGTGA 45226
QY 1372 AGACTTAATCTGAAGCCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1431
DB 45225 AGACTTAATCTGAAGCCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 45182
QY 1432 TCTAGACATCTTAAGGCAAGAACTTCAATGAGCAAGAACTTCAATGAGCAAGAACTTCA 1491
DB 45181 TCTAGACATCTTAAGGCAAGAACTTCAATGAGCAAGAACTTCAATGAGCAAGAACTTCA 45123
QY 1492 AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1550
DB 45122 AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 45063
QY 1551 TCATTAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1610
DB 45062 TCATTAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 45003
QY 1611 CTGACTTAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1670
DB 45002 CTGACTTAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 44943
QY 1671 CAACCAATCTCAAGAGAGGCTTAAGAGCAATGAGCAATCTCAAGAGAGAGAT 1730
DB 44942 CAACCAATCTCAAGAGAGGCTTAAGAGCAATGAGCAATCTCAAGAGAGAGAT 44884
QY 1731 TACCAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1788
DB 44883 TACCAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 44824
QY 1789 AAATCAACACTGTATGAGATACCACTTAATCTCTGCAAGAGATGATTAATTAATTAAT 1848

```

```

DB 44823 AATCAAAACCCGATGATGATCAACCTCACTCCGCAAGAGATGATCAATTAATCAAAAAC 44764
QY 1849 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1908
DB 44763 CAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44704
QY 1909 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1967
DB 44703 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44644
QY 1968 AAATGATGAGCAATTTGATCAAGCAATTTGATCAAGCAATTTGATCAAGCAATTTGAT 2003
DB 44643 AAATGATGAGCAATTTGATCAAGCAATTTGATCAAGCAATTTGATCAAGCAATTTGAT 44584
QY 2004 ----- 2003
DB 44583 AAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44524
QY 2004 -----AATATG 2009
DB 44523 AAATATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 44464
QY 2010 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2069
DB 44463 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44404
QY 2070 ATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2129
DB 44403 ATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44344
QY 2130 GAATGAGAAACCAACATCATATGTTCTCACTTAACAAGTGGGGCTTAAGTGTGAGCA 2189
DB 44343 GAATGAGAAACCAACATCATATGTTCTCACTTAACAAGTGGGGCTTAAGTGTGAGCA 44284
QY 2190 CGAAGGATGAG--AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2248
DB 44283 CGAAGGATGAG--AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44224
QY 2249 CGAAGGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2308
DB 44223 CGAAGGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44166
QY 2309 AATCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2368
DB 44165 AATCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44106
QY 2369 AAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2426
DB 44105 AAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 44048

RESULT 12
US-10-367-094-83
; Sequence 83, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 277616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-367-094-83

Query Match 15.8%; Score 1366.4; DB 17; Length 277616;
Best Local Similarity 76.9%; Pred. No. 36-311;

```


QY 2031 ACCATGATATACACTCAGCCATTAATAAAGATATATGATGACATTCACAGCAATCTAG 2090
 Db 220647 ACAATGAAATTTTACTCGGCAATTAAGAAAGATTAATTAATGACATTTTGCAGGACCCG 220706
 QY 2091 ATGGAATGAGAACCTTATTTCTAAGTGGGGTAATCACTCAGGAATGAGAAACCAACATCAT 2150
 Db 220707 ATGACATGGAACATTAATTTCTAAGTGAAGTAACAGGAATGAGAAACCAACATTTGT 220766
 QY 2151 ATGTTCTCACTTAAGTGGGGCTAAGCTGTGAGACACGAAAGCAT -AGAATGATATA 2209
 Db 220767 ATGCTCTCCCTGATATGTGGAGCTAAGCTATGAGACGCAAGGATTAAGATGATACA 220826
 QY 2210 ATGAATCTGGGAGCTTGAAGGGAAGATGAGAGAGGCGAGATTAAGATCACTACAC 2269
 Db 220827 AGGACCTTTGGGACTTGTGGGAGAGGTGGAGGTGGGCAAGGATTAAGATGATGAAA 220886
 QY 2270 ATGGGTACAGTGTACACTGCTCAGTGTAGTGGTGACCAAAATCTCAGAAATTAACACTA 2329
 Db 220887 TAAGGCAAGTGTATATCTGCTGGGTGATGGGTGACCAAAATCTCAGAAATCACTAC 220946
 QY 2330 AAGAACTTATCCATGAGAGCAACACCACTGTTCCCAAAATCCCAATGAAATTAATAAT 2389
 Db 220947 AAGAACTTATCCATGAGAGCAACACCACTGTTCCCAAAATCTTATGAAAAATTAATA 221006
 QY 2390 AATAATTAATAA 2401
 Db 221007 AAAAAAGCA 221018

RESULT 13

US-09-880-107-3428/c
 ; Sequence 3428, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; PRIOR FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3428
 ; LENGTH: 99014
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
 US-09-880-107-3428

Query Match 15.8%; Score 1358; DB 9; Length 99014;
 Best Local Similarity 78.8%; Pired. No. 1.5e-309;
 Matches 1895; Conservative 0; Mismatches 345; Indels 170; Gaps 16;
 QY 131 AGCATATCAAGAAATATATCCATTTGCAAGTGGTTTCAATACAGAGGGGTGACAGA-T 189
 Db 88976 AGAAAAAATTAAGATATCCACATGCTCAAGTTGTTTCAATACAGAGGATGACAGAGT 88917
 QY 190 AGTTTACATACCAAGTCAATTAATGTGATATCATCAATAAACAGAAATTAACAAAA 249
 Db 88916 GGTTTAATCATGCAAGTTAATCATGTGATATACACCAATTAACAGAAATTAACAAAA 88857
 QY 250 ATCAATGATCATCTCAATATGATGCTGAAAAAGCAATTTGCAAAATTAATTAATTTT 309
 Db 88856 ATCAATGATCATCTCAATATGATGCTGAAAAAGCAATTTGCAAAATTAATTAATTTT 88797
 QY 310 TGATTAACCTTACGCAAAATGACATATGAAGGACATATCTTAATGATTAAGGCCA 369

Db 88796 TGATTAACCTTACGCAAAATGCGCATATCAAGGACATATCTCAATGATTAATTAAGCCA 88737
 QY 370 TATATGAGGAGCCACAGCAAACTATTATCTGAATGGGAAAAATTGAATTTGTTCC 429
 Db 88736 TCTATGAGCAACCCACAGCAAACTATTATCTGAATGGGAAAAATTGAATTTGTTCC 88677
 QY 430 TGAAGACTGGAACAGCAAGGATG-CTACTTTCAACCTTTATTCAAATGATAGG 488
 Db 88676 TGAAGACTGGAACAGCAAGGATGCTCTCACTCTCTTTCAACATGATAGTGG 88617
 QY 489 AAGTTTACCCAGAGCAATCAGCAAGAGAAAGAAATCAAGGGACCCCAATCAATTAAG 548
 Db 88616 AAGTTTACCCAGAGCAATCAGCAAGAGAAAGAAATTAAGGGCATTC-AATGGTAAAG 88558
 QY 549 AGGAAGTCAACTGTCTCTGTTCACTGATGATATGATGATATCTTAAGAAACCTTAAG 608
 Db 88557 AGGAAGTCAACTGTCTGTTGTTGATGATGATATGTTTACCTTAAGAAACCTTAAG 88498
 QY 609 ACTCATCCAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 668
 Db 88497 ACTCATCCAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 88438
 QY 669 TAAATGTACCAAAATCAGTACGCTGTATACCAACAGTACCAAGCTGAGATCAAA 728
 Db 88437 TTAATGTACCAAAATCAGTACGCTGTATATACCAACAGTACCAAGCTGAGATCAAA 88378
 QY 729 TCAAGACTCAAACTTTTACATATGCTTAA-----AATATCTTAAGATATTTCT 782
 Db 88377 TCAAGACTCAAACTTTTACATATGCTTAA-----AATATCTTAAGATATTTCT 88318
 QY 783 TACCCAGAGGATGAGAGCTCTCAAGGAAATCTACAAACAGAGCT---GACATCA 838
 Db 88317 TACCCAGAGGATGAGAGCTCTCAAGGAAATCTACAAACAGAGCTGAGAAAGAGCA 88258
 QY 839 TAGATGACCAAAACAGTGAAGAACATCCATGCTCATGATGATGATGATGATGATGATGAT 898
 Db 88257 TAGATGACCAAAACAGTGAAGAACATCCATGCTCATGATGATGATGATGATGATGATGAT 88198
 QY 899 TGAATGACCAATTTGCAAAAGCAATCTACAGTTCAATGATGATGATGATGATGATGATGAT 958
 Db 88197 TGAATGACCAATTTGCAAAAGCAATCTACAGTTCAATGATGATGATGATGATGATGATGAT 88138
 QY 959 CATCATCATCTTCAAGAGAACTGAGAAAGAAATCAATTTCAATTTGATGAGAAACAAAC 1018
 Db 88137 CATCATCATCTTCAAGAGAACTGAGAAAGAAATCAATTTCAATTTGATGAGAAACAAAC 88085
 QY 1019 CAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1078
 Db 88084 CAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 88033
 QY 1079 GAGGATCATCATTTACCATCTTCAATCTATCTACAGGCTATTAATCAACCAATCAT 1138
 Db 88032 GAGGATCATCATTTACCATCTTCAATCTATCTATTAAGGCTATTAATCAACCAATCAT 87973
 QY 1139 GGCATGACATTAATAGCCAACTGATTTTGAACAAAGCAAAACAAATTAAGTGGGAG 1198
 Db 87972 GGCATGACATTAATAGCCAACTGATTTTGAACAAAGCAAAACAAATTAAGTGGGAG 87913
 QY 1199 AGCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAAAACAAATTAAGTGGGAG 1258
 Db 87912 AGCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAAAACAAATTAAGTGGGAG 87853
 QY 1259 AAGACATTTCTAGTTAAACAAATGCTGATGATTTTGGCAAGCCATGTTGAGAGATGA 1318
 Db 87852 AAGACATTTCTAGTTAAACAAATGCTGATGATTTTGGCAAGCCATGTTGAGAGATGA 87793
 QY 1319 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1378
 Db 87792 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 87734
 QY 1379 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1438

Db	87733	AATTAAGACCTGAACTAATTAATAATTCTGAAAGATTAACATTGAAAAAACCCTTGAGC	87674
Qy	1439	ATTCACTTAGCAAAAGCTTCATGGCCAGAACCCAAAAGTAATGCAACAAAACAAA	1498
Db	87673	ATTGGCTTAGCAAGAACTTCAAGACCAAGAACCCAAATGCAATCAATTAATAACAAAG	87614
Qy	1499	ATTAATGTAGTAGCACTTAATTAACCTTAAGAGCTTTTGGCAGCAAAAACATCAATTGCC	1558
Db	87613	ATTAATATGCTGGGAC----TTAAACTTAAGAGCTTTTGACGAAAAAAGGAACAGTCAGC	87558
Qy	1559	AGAGCAAAACAGACACACCCAGAGTGAGAGAAATCTTCCAAACTTAACATCTGACTTA	1618
Db	87557	AGTGTA--AGAGAACCCACAGAGTGGAGAAAAGCTTCACATCTTAACATCTGACAAA	87500
Qy	1619	GGACTTAATTCGGAATTCACAGAACTCAAAACAATCAGACAGAGAAAGAACAAAT	1678
Db	87499	GGAAATTAATTCAGAAATCTCAATGAAGTCAAAACAATCAGACAGATTAATAACAAACAT	87440
Qy	1679	CCCATGAAAGAGTGGGCTAAGGACATGAATAAGCAATTCCTCAAAAGAAAGTATACAAAG	1738
Db	87439	TCCATCAAAAAGTGGGCTAAGGACATGAATAAGCAATTCCTCAAAAGAAAGTATACAAAG	87380
Qy	1739	GCCAAACAAACGAGAAAAATGCTTAACATCACTAATGATTAGGGAATGTAATCAACAC	1798
Db	87379	-----CATGTGAAAAATGCTCAACATCAACAGATCGGGAAATGCAATCAAAAC	87327
Qy	1799	TGTAAATCGATACCACTTACTCTCGCAAGATGTCATATTTAAAACTTAATAATTA	1858
Db	87326	TGCATATCAATTACTAATCTTACTCTCGCAAGATGTCATATTTAAAACTTAATAATTA	87267
Qy	1859	TAGATGTTGGTGGGTCTGTGTGTATTAAGAACACTTTTACACTGCTGTGGGAATGTAA	1918
Db	87266	TAGATGTTGGTGTGTGTGTGTATTAAGAACAGGGAACTTTTAA--TGTCGTGTGGGAATGTAA	87208
Qy	1919	ACTTGCCCAACCACTATGAGAAAAACAGTGTGAATTTCTTAAGAACTTAAGTAGATGC	1978
Db	87207	ACAAAGTATGGCCACATGAGAAAACAGTGTGAATTTCTTGAAGAACTGAAGTAGACT	87148
Qy	1979	ACCAATTTGATCCAGCAATCCCATTA-----	2003
Db	87147	ACCAATTTGATCCAGCAATCCCATTA-----	2003
Qy	2004	-----	2003
Db	87087	CGAAGAGATACCTGCACACAGCTTATATAGTACACAATTTGCAATTCGAAATACGTGA	87028
Qy	2004	-----	2004
Db	87027	ACCAACCCAAATGCCCATCAATCAACAGTGAATAATTAACCGTGATATCTTAACAA	86968
Qy	2035	TGGAATCAACCTCAGCCATTAATAAGTAATGAATGACATTCACAGCAATCTAGATGG	2094
Db	86967	TGGAATCTACCCACCATTAATAAGTAATGAATGACATTCACAGCAATCTAGATGG	86908
Qy	2095	AATTGAGACCTTATTTCTAAGTGGGTAACCTGAGAAATGGAACCAACATCATATGT	2154
Db	86907	GATTGAGAGCAATTAATTAAGTAAGTAATCTCAGGAATGGAACCAACCTGTATGT	86848
Qy	2155	TTCACACTTAACATGGGGGCTTAAGCTGTAGGACACGAAGGCT--AGATGATATATGA	2213
Db	86847	TTCACACTTAACATGGGGGCTTAAGCTGTAGGATGCAAAAGGCTTAAGATGACACAATGG	86788
Qy	2214	ACTCTGGGAGCT--TGAAGGGAAGGATGGAAGAGAGCGAGAGGTAAAGATCACTACAAATG	2272
Db	86787	ACTTGAAGACTCAGGGGGGAAGAGGTGGAAAGAGATGAGGTATTAAGACTTAACAAATGG	86728
Qy	2273	GGTACAGTGTACACTGCTCAGGTGATGGGTGCACAAATCTCAGAAAT--TACCACTAAA	2331
Db	86727	GGTTCAGTGTACTGTCCCGGGTGAATGGGTGCACAAATCTCAGAAATGCAACTAAA	86688
Qy	2332	GAACTTTATCCATGGAACAAACACACCTTTTCCCAAAATCCCAATGAAATTAATAATTA	2391
Db	86667	GAACTTTATCTATGTAAACAAATTAACACCTGTTTTCCGATTAACCTATGAGAAAAAAACAA	86608

OY		2392	TAAATATAA	2401		
Dd		86607	AAAACAAACA	86598		
		<hr/>				
		RESULT 14 US-10-775-169-149 ; Sequence 149, Application US/10775169 ; Publication No. US20040175743A1 ; GENERAL INFORMATION: ; APPLICANT: Wyeth ; APPLICANT: Burczynski, Michael ; APPLICANT: Twine, Natalie ; APPLICANT: Dornier, Andrew ; APPLICANT: Trepicchio, William ; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo ; FILE REFERENCE: AM101080 (031896-013000) ; CURRENT APPLICATION NUMBER: US/10775,169 ; CURRENT FILING DATE: 2004-02-11 ; NUMBER OF SEQ ID NOS: 5278 ; SOFTWARE: PatentIn version 3.2 ; SEQ ID NO: 149 ; LENGTH: 136284 ; TYPE: DNA ; ORGANISM: Homo sapiens US-10-775-169-149				
		<hr/>				
		Query Match	15.6%;	Score 1346;	DB 17;	Length 136284;
		Best Local Similarity	76.6%;	Pred.	No. 1.3e-306;	
		Matches 1985;	Conservative	0;	Mismatches 390;	Indels 215; Gaps 19;
OY		1	TGTATGAAGCCATTGTCCTTTTAATCCAAAAACCAGAAAAGATATA--CAAAAAAGAA	58		
Dd		83890	TCTATGAAGCAGTGTGCACCTTAATCCAAAAACAGAAAGGGCATATACAAAAAACAA	83949		
OY		59	ACTATAGAACCAGTACCACTGATGAATATACATGCCAGAAATCCCACAAAAATACTAGCTA	118		
Dd		83950	ACGATTAACCGATATCTTGATGAACATATGATGCTGACATCTCTTAACAAAATATTAACCTA	84009		
OY		119	ACCCAATCCAAACGACATATCAAGAAATTAATCCACATTTGTCAAATGGGTTTCATACCAG	178		
Dd		84010	ACTGAATCCAAACGCAATTCAMAAAAATATCACACATATCAAGTGGGTTTCATACCAG	84068		
OY		179	GGGTGCGAG-AATAGTTAAATCATACAGACATTAATGTGATPACATCAATAAACAGAA	237		
Dd		84070	GGAATGCGAATGGTTTTAAATAACAAAGTAATTAACSTGATACCCACATPAAGCAGAA	84128		
OY		238	TTAAAAAC-AAAAATCAATGATCATCTCAATAGATGCTGA AAAAGCATTTGACAAATC	296		
Dd		84130	TTAAAAACAAAATAATCATGATCATCTCAATGATGCGA AAABAGATTC AACAAAATC	84189		
OY		297	TAAcATTCTTTATGATTAAAACTTACGAAAATCGACATNAGAAAGCATACCTT---	353		
Dd		84190	CAGCATCCCTTTATGATTTAAAACTTCACAAAAATCAGCATVCAAGAAACACACCTTGT	84249		
OY		354	-----AATGTAAATAAAGCATATATGAGAGSACCCACACAAACATTATATCGAATG	405		
Dd		84250	TGTATTAATAATGTATTAATAAGCCAACTATGACAAACCCACAGTCACATTAATACAAATG	84309		
OY		406	GGGAAAAGTTGAAAACATTTGTCC-----CTGAGAACTGGANAC	442		
Dd		84310	GGGAAAAGTTGAAAAGCATTTCCCTCTGAGAAAGCTGAAGAAGCATTTCCCTGAGAACTGGANAC	84366		
OY		443	AAGACAAGATG-CYACTTTACCACTTTCTATTCAACATATATGATGGAAGTTTATGCCAG	501		
Dd		84370	AAGACAAGATGCCCCACTTCACACACTCTTCTTCAACAACATATGGAAGTCCATGCCAG	84429		
OY		502	AGCAATACAGACAGAGAAAGAAATTC AAGSGCACCCAAATCAATTAATAAGAGAGTCAAATCT	561		
Dd		84430	AGCAATACAGACAGAGAAAGAAATTAATAAGAGATTC AATTGGTTAAAGAGAGTCAAATCT	84489		
OY		562	GTCCCTGTTTCACTGATGATATGATTTGATATTAACCTTAAGAAAACCTTAAGAATCTCAAGAA	621		

Db 84490 GTCACTGTTGCTGATGATATGATTTGATCTGAAAACCTTAAAGATCTCCGAGAA 84549
Qy 632 GCTCTAAGACGTATACATTAATTCGTAAAGTTTCAGATACAACTAAATGATACAA 681
Db 84550 GCTCTAAGATGATTAAGAAATCCAAAGATTTCTGATACAAAGCTCATGTATACAA 84609
Qy 682 ATCAGTACGCTGTATACACCAACAGTACCAAGCTGAGAAATCAATCAAGAACTCAA 741
Db 84610 ATCAGTACGCTGTATACACCAAGATGATGACCAAGGAGAAATCAATCA--GACTCAA 84667
Qy 742 CACTTTTCAATAGCTGTAA-----AAAATCTTAAGAAATTTTCCCAAGAG 793
Db 84668 CCGTTTCAATAGCTGTCAAAAATAAATAATCTAGCAATATATACCTAACCAAGAG 84727
Qy 794 GTGAAGACCTTACAGAGAAATCAAAACACACT---GACTCATAGTACACA 849
Db 84728 TCAAAAGACCTGTACAGAGAAATCAAAACACCTGTAAGAAATCAAGATGACACA 84787
Qy 850 AACAGTGGAAACACATCCCATGCTCATGAGTGGATGAGATCAATATTTGAAATGACC 909
Db 84788 AACAAATGAAACATCCCATGCTCATGAGTGGATGAGATCAATATTTGAAATGACC 84847
Qy 910 ATATTGCCAAAAGCAATCTACAAAGTTCAATGCAATTTCCACCAAAATATCATCATTC 969
Db 84848 ACATGCCAAAAGCAATCTACAAATTCATATGCAATCCCATCAAAATACCAATCATTC 84907
Qy 970 TTCAAGAACTAGAAAAAACAATTTCTAAATTCATATGAGAACAAACCAAAAAA 1029
Db 84908 TTCAAGAAATTAAG--AAAAACAATTTCTGAATTTATATGAGAC-----CAA 84951
Qy 1030 AAAAAAACCCGATAGCCAAAGCAAGCTTAGCAAAAAGCAAAATCTGGAGGCATCACA 1089
Db 84952 AAGAGAGCCAGATAGCCAAAGCAAGCTTAGCAAAAAGCAAAATCTGGAGGCATCACA 85011
Qy 1090 TTACCATCTTCAAACTATATCAAGGCTATATATCAACAAACATCATGACCTGACAT 1149
Db 85012 CTACCTGATTTCAAACTCTATATATAGGCCATGTCAGAAAGCAATGTCGTCTGTAT 85071
Qy 1150 AAAATAGGCAATGACCAAT-----GGAAGAGAGAGAAATCCGAATTAATAGC 1201
Db 85072 AAAAAATAGCAATGAGCACTATACAGAAATGAGAAAGAAATGAGAAACCCGAATTAATC 85131
Qy 1202 CAATAATATATATAGCCAACTGATTTTTCAGAAAGCAAAACAAATTAATGAGGAAAG 1261
Db 85132 CAATAATATATAGCCAACTGATTTTTCAGAAAGCAAAACAAATTAATGAGGAAAG 85190
Qy 1282 ACATTTAGTAAACAAATGCTGAGATTTTGGCAAGCCACATGTGAAAGATGAAC 1321
Db 85191 ACACCTTTTCAACAAATGCTGAGATTTTGGCTAGCCACAGGTAGAGATGAAC 85250
Qy 1332 TGGATCCCTGCTCTCTCACTTAATACAAAAATGATACAAAGATGATCAAAAGCTTAAT 1381
Db 85251 TGGAT--CCTCATCTCTCTCTCTTAATACAAATCAACCTCAAGATGATTAAGCTTAAC 85309
Qy 1382 CTGAGACCTTAACCAATTAATAATTTCTAGAAGATTAATCAAGAAATGCTTCTAGACAT 1441
Db 85310 CTGAAGCTGAAACCTACAAACATCTCTAGAAATTAATTTGAAAGAAACCTTCTAGACAT 85369
Qy 1442 CACTTAGGCAAGACTTCAATGGCCAAAGAACCAAAAGTAATGCAAAAAACAAATA 1501
Db 85370 GACTTAGGCAAGACTTCAATGACCAAGAACCAAAACAAATGCAATTAATAAAGATA 85429
Qy 1502 AATAGATAGACTTAATTAATTAATAAGCTTTTGGCGACCAAAACAAATCATTAAGACA 1561
Db 85430 AATAGATAGACTTAATTAATTAATAAGCTTTTGGCGACCAAAACAAATCATTAAGACA 85488
Qy 1562 GCAAAACAGCAACCCAGAGTGAAGAAATCTTCAAACTAGACATCTGACTAAGA 1621
Db 85489 GTAAACAGCAACCCAGAGTGAAGAAATCTTCAAACTAGACATCTGACTAAGA 85548
Qy 1622 CTAAATATCCGAATTCACAGAGAACTCAAACAATGACAGAGAAAGAAACAAATCCC 1681

Db 85549 CTAAATCCAGAACTACAAAGCACTCAAAACAAATCAAGTAAGAAAA--AAACATCCC 85606
Qy 1682 ATGAAGAGTGGGCTTAAGACATGATGATGACATTTCTCAAAAGATATACAAATGACC 1741
Db 85607 ATCAATATGTTGGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 85666
Qy 1742 AACAAAC--AGAAAAAATGCTTAAATCACTAATGATTAAGAAATGTAATCAACT 1799
Db 85667 AACAAACATATGAAAAAATGTTCAATCACTAATGATGATGATGATGATGATGATGAT 85726
Qy 1800 GTAAATGCAATCACTTCTCTCTGAGAAATGATGATTAATTAATAATCTTAATAAT 1859
Db 85727 ACATGTCGATTAATCTTCTCTCTGAGAAATGATGATTAATTAATAATCTTAATAAT 85786
Qy 1860 AGATGTTGGTGGGCTCTGCTGATTAAGAAACATTTTACATGCTGCTGAGAAATGTA 1919
Db 85787 AATGTTGGTGGGCTCTGCTGATTAAGAAACATTTTACATGCTGCTGAGAAATGTA 85846
Qy 1920 CTTGCGCAACCACTATGAAAAACAGTGGAAATTTCTTAAGAACTTAAGATGATGAT 1979
Db 85847 CTAGTACAGCATTAATGAAAAACAGTGGAAATTTCTTAAGAACTTAAGATGATGAT 85906
Qy 1980 CCAATTTGATCCAGCAATCCATTT----- 2002
Db 85907 TGGTTGATCCAGCAATCCATTTCTGCTGATGATGATGATGATGATGATGATGATGAT 85966
Qy 2003 ----- 2002
Db 85967 AAAAAAGATGCTTGCATGACAGGTTTATAGACCAATTCATCAATTCGAAATGCTGG 86026
Qy 2003 -----AATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2029
Db 86027 AATCAACCAAAATCCCATCATCAATGATGATGATGATGATGATGATGATGATGATGAT 86086
Qy 2030 TACATGAAATTAATCACTGACCCATTAATAAAGATTAATGATGATTAATCAAGCAAT 2089
Db 86087 TATGATGAAATTAATCACTGACCCATTAATAAAGATTAATGATGATTAATCAAGCAAT 86146
Qy 2090 GATGAAATGGAAGACCTTATTTCTAAGTGGGTAATCAAGAAATGGAAGAAACCAATCA 2149
Db 86147 AATGAACTGAGACATTAATTTCTAAGTGGGTAATCAAGAAATGGAAGAAACCAATCA 86206
Qy 2150 TATGTTCTCACTTAAGAGTGGGCTTAAGTGGGTAATCAAGAAATGGAAGAAACCAAT 2208
Db 86207 TATGTTCTCACTTAAGAGTGGGCTTAAGTGGGTAATCAAGAAATGGAAGAAACCAAT 86266
Qy 2209 AATGAACCTGAGGACTTAAGAGGAAAGATGAAAGAGGCGAGGATTAAGAAATCAAC 2268
Db 86267 AATGAACCTGAGGACTTAAGAGGAAAGATGAAAGAGGCGAGGATTAAGAAATCAAC 86326
Qy 2269 AATGGTACAGTGTACATGCTCAGGTGATGGGTGCAACCAATCTCAGAAATTAACACT 2328
Db 86327 AATGGTACAGTGTATCTGCTCAGGTGATGGGTGCAACCAATCTCAGAAATTAACACT 86386
Qy 2329 AAGAACTTATCCATGAGAAAGCAACCACTGTTCCCAAAATCCCAATGAATTAATA 2388
Db 86387 AAGAACTTATCCATGAGAAAGCAACCACTGTTCCCAAAATCCCAATGAATTAATA 86446
Qy 2389 TAATTAAT 2398
Db 86447 GATGACACTT 86456

RESULT 15

US-09-795-668-1
; Sequence 1, Application US/09795668
; Patent No. US2002045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinhordottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001

Query Match	15.2%	Score	1311.6	DB	9	Length	1503841
Best Local Similarity	76.3%	Pred. No.	8.4e-298				
Matches 1937; Conservative	3	Mismatches	432	Indels	168	Gaps	20

Qy	1	IGTATGAAGCCCATCTCATTAAATACAAAACGAGAAGATAT-ACAAAAAGAAA	59
Db	545389	TCTCAAAAGCCAGTATCACCCCTAATGCCAAGCGAGAAAGCATACAAAABACGAA	546048
Qy	60	CTATAGACCAGTACCATGTATGAATTTATCATGCAGAAATCCCACAATAACTACTTA	119
Db	546049	ACTACAGTCAGTATCTCGATGAACCTATGATCGAAAATCTCTCAATTAATCTTAGGAA	546108
Qy	120	CCCATCCAACGCGATATCAAGAGATATCCACCATTTGCACAGTGGGTTTCATACAGG	179
Db	546109	TTGATATCCAAGTATATCAAABAAGTATATCCACCATGATCAGTAGTTCACATAGG	546168
Qy	180	GGTGCA-GGATAGGTTTAACATACACAAGTCATAAATGTGATCATCATTAACGAAT	238
Db	546169	GATGACAGGATGGTTTTAAACATAGGCMAAGTCATAAATGTGATCACACACTTAAACGAAT	546228
Qy	239	TAAAAACAAAATCATCATGATCATCTCATPATAGTGTGAAAAAGCATTTGACAAAAATCTA	298
Db	546229	TACAACAGGAATCATCATGATTTGTCTCATATGACACAGAAAAACATTTGACAAAATCCA	546288
Qy	299	ACATTTCTTTATGATTAANAACCTTCAGCAAATCGACATAGAAAAGCATACTTAATGT	358
Db	546289	ATTCCTCTTTTGACTTAANAACCTCAGCAAATCTGCATATGAAAGGGTCTATCTTAAGT	546348

OY	359	ATAATAACCCATATATGACGGAGCCCAACGCAAACTTTATACGATATGGGGGAAAAAGTTGAA	418
Db	546349	AATATAAACCATCTATGCGCAGGCCCAACGCCAACTTTATATCTGAGAGGAAAAAGTGGAA	546400
OY	419	AACATTTGCTCCGAGAACTGGAAACAAGACMAAGATGCTACTTTCACCACTTCTATTCAAC	478
Db	546409	AGCATTTCCCATTAAGAACTGGAAACAAG-----GATGCCCACTTTCACCACTTCTATTCAAC	546460
OY	479	ATAGATAGTGAAGTTTATGCGCAGACCATACAGAAAGAAAAATCAAGGGCACCCAA	538
Db	546465	ATCGTACAGGAATTTAGCCAGACCAATCAACAAGACAAAGAAATTAAGACATCCAA	546520
OY	539	ATCATATAAAGAGAAAGTCAAACTGCTCCCTGTTCACTGATGATATGATTGTATCTAGAA	598
Db	546525	AATGTTAAGAAGACATCAAAATTTGTTGGCTGTTTGACAGATGATATGATGTATACCTAG-A	546580
OY	599	AACTCTAAGACTCATCCAGAAAGCTCTGAACCTGATCATTAATTCAGTAAAGTTTCA	658
Db	546584	AACTCTAAGATTTCCACCAAAAAGCTCTGAACCTGTTAAAGAAATTCAGCAAAAGTTTCA	546640
OY	659	GGATCACAAACTAAATGTACACAAATCAGTAGCAGTCTATACCAACAAGTGACCAAGCT	718
Db	546644	GGATTCAAAATTTAATGTAAA-AAATTAAGTACTGTCTATACCTATACAGCAACAAGCT	546700
OY	719	GAGAAATCAAAATCAAGAACTCAAAACCTTTTCAATAGCTGTAAA-----AAAATCTTAA	773
Db	546703	GAGAAATCAAAATCAAGAACTCAAAACCTTTTCAATATCAGCAAAAATAATTAATACRTAG	546760
OY	774	GAATATCTTATCCCAAGAGGTGAAGAGCTCTACAGGAAAACTACAAAACACACT--	831
Db	546763	GAATATCTTAAACAGAGATGGAAGAGCTCTTACAGGAAAACTACAAAACACACTCTGA	546820
OY	832	--GACATCATAGATGACACAAACAAGTGGAAAACATCCCATGCTCATGATGGTGTAGAA	889
Db	546823	AAGAAATCATAGACACACAAACAATGGAAAACATCCCATGCTCATGATGGGT-GAA	546880
OY	890	TCATATTTGTGAAAAATGACCATATTGCCAAAAGCAATCTTACAGTTCAATGCAATTCCTCA	949
Db	546882	TTAATATTGTGAAAAATGACCATATTGCCAAAAGCAATCTTACATGCAATTCCTCT	546940
OY	950	CCAAAATTCATCATCTTCTTCAACAGAACTAGAAAAAAACAATCTTAAATTCATATGG	1009
Db	546942	TCAAAATTCATCATCTTCTTCAATATATG-AACAAAACAATCTTAAATTCATATGG	547000
OY	1010	AACAAACAACCAAAAAAATAAAAAAACCCTGATGCCAAGCAAGACTTTAGCCAAAAGA	1069
Db	547001	AAC-----AAAAAAGAGCTGATAGCAAAACAAGACTTAAAG-CAAAAG	547045
OY	1070	ACAAATCTGGAGCATCATTTACCCATCTTCAAACTATACTTACAAAGCTATATATACCA	1129
Db	547046	ACAAATCTGGAGCATCATCTTATCTTCAATATATATTAATTTAAGGCAATATAGTACCA	547105
OY	1130	AAACATCATGCGCATGACATATAAACTAGGACATAGACCAATGAAAAAGAAGAGAAATC	1189
Db	547106	AAACGATCTGTATCTGTATATAAAATAGTGATAGACCAACAAATGGAACAGATATGAGAAC	547165
OY	1190	CAGAAATTAAGCCCAATATATTAATAGCCAACTGATTTTGGCAAGCAAAACAATAA	1249
Db	547166	CAGAAATTAAGCCCAATATCTTAACGCCCACTGATATTTCAACAAAGCAAAACAATAA-C-AA	547220
OY	1250	AGTGGGAAAAAGACATTTAGTTTAACAATGGTGTGAGATTAATTGGCAAGCCACATGTG	1309
Db	547224	AGTGGGAAAAAGACCTCTTATTCACAAATGGTGTGAGATATTTGGCAAGCCACATGTGA	547280
OY	1310	GAAAGATGAACCTGGATCCCTGTGCTCTCATCTTAAATACAAAATTTGATACAAGATGATC	1369
Db	547284	GAGGATGAACCTGGAT-CTTTCATCTTATCTTATGCAAAAATCACTCAAGATGATC	547340
OY	1370	AAAAGACTTAATCTGAGACTTAAAAACATTAATAATTTCTAGAAATTAACATCGAAAAATG	1429
Db	547343	ATAATCTTAAATCTTAAAGATGTGAACCAATAAAGATTTTGAAGAGATTAACATCGAAAAAAC	547400

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:16:09 ; Search time 22711 Seconds
(without alignments)
13833.973 Million cell updates/sec

Title: US-09-867-570-3
Perfect score: 8622
Sequence: 1 tgcataagacatgtcact.....gtccccaagccctacc 8622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_esc3: *
4: gb_esc4: *
5: gb_esc5: *
6: gb_esc6: *
7: gb_esc7: *
8: gb_esc8: *
9: gb_esc9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020.8	11.8	2703	8 AF101616	AF101616 AF101616
2	877.6	10.2	2735	3 HSM804825	AL833512 Homo sapi
3	814.6	9.4	5133	3 AY539945	AY539945 Rattus no
4	801.8	9.3	3654	3 AY387074	AY387074 Rattus no
5	801.8	9.3	4071	3 AY539883	AY539883 Rattus no
6	790.6	9.2	4356	3 AY539886	AY539886 Rattus no
7	770	8.9	2488	3 BC038425	BC038425 Homo sapi
8	769.2	8.9	2231	3 HSM804900	AL833512 Homo sapi
9	749.2	8.7	6966	3 AY387112	AY387112 Rattus no
10	747.8	8.7	3429	3 AY539894	AY539894 Rattus no
11	744.8	8.6	8575	3 CR627265	CR627265 Homo sapi
12	720.4	8.4	3232	3 AK038045	AK038045 Mus muscu
13	708.2	8.2	889	6 CA455045	CA455045 AGENCOURT
14	702.8	8.2	2776	3 BC028293	BC028293 Homo sapi
15	683.4	7.9	871	7 CN839125	CN839125 AGENCOURT
16	682.2	7.8	3315	3 AY387051	AY387051 Rattus no
17	663.2	7.7	863	7 CN835944	CN835944 AGENCOURT
18	655.8	7.6	2354	3 AK052675	AK052675 Mus muscu
19	645	7.5	5064	3 AY539942	AY539942 Rattus no
20	631.6	7.3	3309	3 AY387094	AY387094 Rattus no
21	627.2	7.3	1821	8 AF101873	AF101873 AF101873
22	627	7.3	1897	3 AY325175	AY325175 Rattus no
23	620.2	7.2	3523	3 BC035112	BC035112 Homo sapi
24	610.6	7.1	3222	3 AY539902	AY539902 Rattus no

25	607.4	7.0	851	7 CN839827	CN839827 AGENCOURT
26	606	7.0	835	7 CN843633	CN843633 AGENCOURT
27	604.4	7.0	764	4 BG198766	BG198766 RST18035
28	603.2	7.0	827	7 CN831885	CN831885 AGENCOURT
29	601.8	7.0	3662	3 HSM804836	AL833523 Homo sapi
30	598.2	6.9	924	7 CN839634	CN839634 AGENCOURT
31	596.2	6.9	3579	3 AK048255	AK048255 Mus muscu
32	589.4	6.8	5145	3 AY539881	AY539881 Rattus no
33	585.8	6.8	688	2 BE439409	BE439409 HTM1-025F
34	585.4	6.8	899	7 CN843534	CN843534 AGENCOURT
35	582	6.8	835	7 CN834097	CN834097 AGENCOURT
36	581.8	6.7	826	7 CN839711	CN839711 AGENCOURT
37	567.4	6.6	2855	3 AK080274	AK080274 Mus muscu
38	560.2	6.5	699	8 AQ286739	AQ286739 RPT11-89
39	557.4	6.5	825	7 CN835770	CN835770 AGENCOURT
40	551.6	6.4	1076	8 B2611142	B2611142 WHAC163TR
41	548.2	6.4	787	4 BG208126	BG208126 RST27616
42	546	6.3	855	7 CN839609	CN839609 AGENCOURT
43	544.6	6.3	825	7 CN843258	CN843258 AGENCOURT
44	544.2	6.3	9860	3 HSM80687	BK640640 Homo sapi
45	540	6.3	4583	3 CR627206	CR627206 Homo sapi

ALIGNMENTS

RESULT 1	AF101616	2703 bp	DNA	linear	GSS 06-NOV-2000
LOCUS	AF101616				
DEFINITION	AF101616 Human Homo sapiens genomic clone pTMB233.328P6, genomic survey sequence.				
ACCESSION	AF101616				
VERSION	AF101616.1	GI:4249719			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Beppler, G., O'Brian, K.C., Kim, Y.C., Schreiber, G. and Pitterle, D.M.				
TITLE	A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH1A metastasis suppressor region				
JOURNAL	Genomics 55 (2), 164-175 (1999)				
MEDLINE	99134294				
PUBMED	9933563				
COMMENT	Contact: Beppler G Medicine and Radiology Duke University Medical Center Box 2610, MSRB, Room 117, Durham, NC 27710, USA part of a 1.4 megabase contig including the LOH1A metastasis suppressor region Bin B Class: unknown.				
FEATURES	Location/Qualifiers				
source	1..2703 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /map="11p15.5" /clone="pTMB233.328P6" /clone_id="Human"				
ORIGIN					
Query Match	11.8%; Score 1020.8; DB 8; Length 2703;				
Best Local Similarity	73.6%; Pred. No. 1.8e-216;				
Matches 1486; Conservatave	0; Mismatches 497; Indels 37; Gaps 13;				
Qy	1 TGTATGAGCGCAATGTCATTTAATACCAAGAGAGATATA--CAAAAAGAAA 58				
Db	91 TTTATGAGCGCAATGTCATTTAATACCAAGAGAGATATA--CAAAAAGAAA 150				
Qy	59 ACRATGACCAAGTACCAATGATATATACATGACGAATCCCAAAATATCTGCTA 118				

Db 151 ATGTTACACCAATATCTTGATGAACATGATGCAAAAATCTCATATAAATCTGGCA 210
 Qy 119 ACCCAATCCACAGATATCAAGAGATATCCCATTTGCAAGTGGTTTCAACAG 178
 Db 211 ACTGATTCAGAGACATCAAAAAGCTTATCCACCATGATCAAGGGCTTCATCCG 270
 Qy 179 GGGTGC -AGSANTGGTTAATATCAACAAGTCAATTAATGATGATGATCAATAAGAA 237
 Db 271 CGATGAAAGGCTGGTTCAACATGCAAAAATCATTAATCAATTCAGCATTAAGCAAA 330
 Qy 238 TTAATAAACAATAATCATGATCATCTCAATAGATCTGAAAAAGCATTTGACAAAATCT 297
 Db 331 CCAAGACACAAACCATATGATTTATCTCATATGATGACAAAAAGGCTTTGACAAAATTC 390
 Qy 298 AACATTTCTTATGATTAATAACCTTGAGCAAAATGACATGAAAGACATACCTTAATG 357
 Db 391 AACATGCTTACGTCTAAAATCTCATATAATTTAGGTATGATGGAATATCTCAAAA 450
 Qy 358 TAAATAAAGCCATATATGACGACCAAGCAAACTTAATATGAAATGAGGAAATGGA 417
 Db 451 TAAATAGAGCTATCTATGACAAACCAAGCCAAATCATCTGATGAGCAAAAATCTG 510
 Qy 418 AACATTTGCTGGAATCTGGAACAAGCAAGATG -CTACTTTCACCACTTCTATTC 476
 Db 511 AAGCATTCCTTTGAAAACTGGCAACAAGGCAAGGATGCCCTCTCTACCACTCTATTC 570
 Qy 477 ACATATGATGAGATTTTAGCCAGACATCAGACAGAGAAAGAAATCAAGGCAACC 536
 Db 571 ACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
 Qy 537 AAATCAATTAAGAGAGTCAAACTGCTGTTCTGATGATGATGATGATGATGATGATG 596
 Db 631 AATTAAGAAAAAGG -AGTGAATATGCTCTGTTGACATGATGATGATGATGATGATG 689
 Qy 597 AAAACCTTAAGACTCATCAGAAAGCTCTAGAACTGATATACATAATTCAGTAAATG 656
 Db 690 AAAACCTTAAGACTCATCAGAAAGCTCTAGAACTGATATGATGATGATGATGATG 749
 Qy 657 CAGATATCAATTAATGATACAAATCAGTACGACTGCTATACACCAAGTGAAG 716
 Db 750 CAGATATTAATTAAGTGTGCAAAAATCAAGTATTCCTATGACCAATTAAGCAAAA 809
 Qy 717 CTGAATCAATTAAGTGAAGTCAAACTTTTACATAGCTGAAA -AAAAATCT 770
 Db 810 CAGAGAGCCAAATCATGATGAACTCCATTCACAAATGCTTCAAGAGAAATTAATTC 869
 Qy 771 TAAAGATATTTCTTACCAAGAGAGTGAAGGACTCTTACAGAGAAATCTACAAAACAG 830
 Db 870 TAGGATTCACACTTACAGAGAGATGTAAGGACCTTTCAAGAGAAATCTACAAACAG 929
 Qy 831 ---TGACATCATATGATGACAAACAGTGAAGACATCCATGCTCATGATGAGTA 886
 Db 930 TCAATGAATTAAGAGATGATACAAACATTAAGAAATTCATGCTTATGATGAGTA 989
 Qy 887 GAAATCAATTTGTAAGAAATGACATATGCGAAAGCAATCTACAAAGTGAAGTGAATTC 946
 Db 990 GAAATC -ATATGCGTGAAGAAATGCTATGCTGCGAAAGTAAATTAATTAAGTCAATTC 1048
 Qy 947 CCACCAAAATATCATCATATCTTCAAGAGATGAGAAATTAATTAATTAATTAATTA 1006
 Db 1049 CCATCAAGCTTACCAAGATCTTTCTTACAGAACT -GAAAAAACTACTTTAAAGTTCTATA 1107
 Qy 1007 TGGAAACAAACCAAAAAAAGAAAAAAGGCGATAGCGCAAGCAAGCAAGCTTAAGCAAA 1066
 Db 1108 TGG-----AATCAAAAAAGAGCCCAATCACCAAGGCAATCTTAAGCCAAA 1153
 Qy 1067 AGAACAATCTGAGAGGATCACTTACCAATCTTCAAACTATATCTTCAAGGCTATATCA 1126
 Db 1154 AGAACAAGCTGAGAGGATCACTGCTACCTTCAAACTATATCTTCAAAAGCTTACGTTAA 1213
 Qy 1127 CCAAAACATCATGAGCTGATCATTAATACTGAGCATATGAGCAATGAGAAAGAGAGA 1186
 Db 1214 CCAAAACAGCATGATCTGTTACCAAAACAGAGATCTAGCAATGAGAAAGAGAGAGC 1273

Qy 1187 ATCCAAAAATTAAGCAAAATTAATTAAGCAAGTATTTTGAAGCAAGCAAAACA 1246
 Db 1274 CCTCAAAAAATTAAGCGGATATCTACAACTATCCATCTTGAACAACTGAGAGAAACA 1333
 Qy 1247 TAAAGTGGGAAAGACATTTAGTTAACAATGATGCTGAGATTAATTTGCAAGCCAT 1306
 Db 1334 AGCAATGGGAAAGATTCCTATTTAAATATGATGCTGGGAAAGATCGGTACGATAT 1393
 Qy 1307 GTGGAAGATGAATCGATATCCCTGCTCTCTCATTAATACAAAATTTGATACAGATG 1366
 Db 1394 GTAGAAAGCTGAATCGATCC -TTCCCTTACACCTTATACAAAATTAATTAAGATG 1451
 Qy 1367 ATCAAGGCTTAAATCTGAGAGCTTAAACCAATTAATTTCTAGAAATTAATCAAGAAA 1426
 Db 1452 ATTAAGACTTAAACATTAAGACTTAAACCAATTAATTTCTAGAAATTAATTAAGAT 1511
 Qy 1427 ATGCTTCTAGACATTCATTAAGCAAGATTCATGCGCAAGAACCAAAAGTAATGCA 1486
 Db 1512 ACCATTCAGACATATGAGCATGGGCAAGGATTCATGCTTAAACCAACCAAAAGTGA 1571
 Qy 1487 ACAAAAACAAAATTAATATAGATAGGACTTAATTAATTAATTAATTAATTAATTA 1546
 Db 1572 ACAAAAACAAAATTAATAGGACAAACGGGATCTAAATTAATTAATTAATTAATTA 1631
 Qy 1547 ACATATTAAGAGAGCAACAGACACCAAGGAGTGAAGAAATGTCACAAACTTA 1606
 Db 1632 GAAATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651
 Qy 1607 GCATCTGATAGACTAATATATCCGAAATCCACAAAGAACTCAAAATGACAGAAAG 1666
 Db 1692 TCATCTGACAAAGGCTAATATTCAGAAATCTATGATGATGATGATGATGATGATG 1750
 Qy 1667 AAAGCAACATCCATGAAAGAGTGGTAAAGACATGATGATGATGATGATGATGATG 1726
 Db 1751 AAAGCAACATCCATGAAAGAGTGGTAAAGACATGATGATGATGATGATGATGATG 1810
 Qy 1727 GATATCAATGAGCAACAA -ACAGAAAAATGCTTAACATCACTAATGATTAAGGA 1784
 Db 1811 GACATTTATGATATCAAAAAACATGAAGAAATCTCATCTACATGACATGAGAA 1870
 Qy 1785 ATGTAATCAACACTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1844
 Db 1871 ATGCAATTAACCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1930
 Qy 1845 AAATCTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1903
 Db 1931 AAGTCAGAAACAAACAGTGTGAGAGATGAGAGAAACAGAAACAACTTTTACACTG 1990
 Qy 1904 CTGCTGGGAATGTAATCTTGGCCCAACCACTATGAGAAACAGTGTGAAATTTCTTAAGA 1963
 Db 1991 TTGGTGGAGCTGTAATCTGATTAACCACTTGGGAAAGTCAAGTGTGCAATTCCTAGGA 2050
 Qy 1964 ACTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2003
 Db 2051 TCTAGAACTGAATTAATCAATTTGACCCAGCATCCCATTA 2090

RESULT 2
 HSM804825 2735 bp mRNA linear HTC 03-AUG-2004
 LOCUS HSM804825
 DEFINITION Homo sapiens mRNA, cDNA DKFZp686B0737 (from clone DKFZp686B0737).
 ACCESSION AL833512
 VERSION AL833512.1 GI:21734156
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2735)
 AUTHORS Bahr, A., Lauber, D., Mewes, H.W., Weil, B., Amidi, C., Osanger, A.,
 Pohl, G., Han, M., and Wiemann, S.
 CONSRM The German cDNA Consortium

QY	1733	CAAAATGGCAACA--AACAAGGAAAAAATGCTTAACATCATGAATGATTAAGGAAATGAAA	1790
Db	2232	CAAAATGGCAAGAAGGTATATATGTAAGATTGCTTCACATGCTTAATATATACAGAAATGCAA	2291
QY	1791	ATCAACACTGTAATATGCGATACCACTTACTCTCTGCAGAAATGTCATTAATTTAAAAATCT	1850
Db	2292	ATCAAAATCTACAAATGATGATATCATCTTATCCCAAGTTAAATTAAGGCTTTTCCAAAAGACC	2351
QY	1851	AAAAATAATATGATGTTGGTGGGTCTGTGTGTGATTAAGGAACACTTTTACATGCTGTGGG	1910
Db	2352	GGCAATATACACTTGGTGTGGAAGATGTGAGAAAAAGAAACCCTTATTAATGTATGGGG	2411
QY	1911	GAATTAACACTTGGCCCAACCACTATATGAGAAAAACAGTGTGAAATTTCTTAAGAACTAAAA	1970
Db	2412	GAATTAATTAATTAATGTCAGACCACTACGAGGAGAACAGCTTTGGAGATTCTCAGAAAACTAAAA	2471
QY	1971	GTAGATGACACCATTTGATCCAGCAATCCCATTT	2002
Db	2472	ATAGAGTACCATATGATATCCAGCAATCCCACT	2503
RESULT 3			
AY5339945		5133 bp	mRNA linear HTC 21-MAR-2004
LOCUS	AY5339945		
DEFINITION	Rattus norvegicus LRRGT00194 mRNA, complete cds.		
ACCESSION	AY5339945		
VERSION	AY5339945.1	GI:45478227	
KEYWORDS	HTC.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 5133)		
AUTHORS	Xu,C.S., Zhang,L., Chang,C.F., Han,H.P., Wang,G.P., Chai,L.Q., Yuan,J.Y., Yang,K.J., Zhao,L.F., Ma,H., Wang,L., Wang,S.F., Xing,X.K., Shen,G.M., Shi,J.B., Rahman,S., Wang,Q.N. and Zhang,J.B.		
TITLE	Liver regeneration after PH		
JOURNAL	Unpublished		
REFERENCES	2 (bases 1 to 5133)		
AUTHORS	Xu,C.S., Zhang,L., Chang,C.F., Han,H.P., Wang,G.P., Chai,L.Q., Yuan,J.Y., Yang,K.J., Zhao,L.F., Ma,H., Wang,L., Wang,S.F., Xing,X.K., Shen,G.M., Shi,J.B., Rahman,S., Wang,Q.N. and Zhang,J.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-FEB-2004) Heman Bioengineering Key Lab, Heman Normal University, NO.148 Jianshe Road, Xinxiang City, Heman 453002, P.R.China		
FEATURES			
source	Location/Qualifiers		
	1..5133		
	/organism="Rattus norvegicus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10116"		
	1..5133		
	/note="Liver regeneration-related"		
	/codon_start=1		
	/product="LRRGT00194"		
	/protein_id="AA566285.1"		
	/db_xref="GI:45478228"		
	/translation="MSLVIKKSPQITTEINDKAFGRQTVTVTVTSLDGCLPRESEPT RROLQAPVSMHFLASLISGFLVYVIGIDPQRKYKASGLPWGSAOROPCORHNR TLKEDRINSLSHPNPGVGRASPRQSLSGNQKLLIVHSPRPEENTKHLIPWCTSGH BAFGGGTIGIPCCCHREPVGSTRPANSLSGTTELAOSNRNDIMAGKRNNSNRNDR ODVMSSEBPSPTKTMMEYVPEKODIVSSYLIMLIDPKRYKSLISLETQENIKD QVEAYRBSQSKLKAFQENTIKOLKEIMQEBALIKEMMETTLDINOKKROGAVDTS FTNRIOEMERISGAEDSIEIIDSTVKNVRRKLLVNIQIEDISMRSLRITGIEI ESSEDOLGPVNI FKKIILENPNLKKEIPIGIOEAYRTPRLDQKRTSHIIVKPEPI NAONERILIKAREGQVLYKGRPIRIPDSPEPMKARASMTDIVITLREHKOPRLR LYPALSLINIDETKI PHDKTFPEVLSLNLALOR INKGAKOHRASVLAENHEHLD KDRHLYKGMKTI FOANGCKOKCAOVALLISKNIPOLKVIKKEGFIPIKIKING DELSLINTIYAPTRAPTYKETELLKUKTHIAPIHTIVDENTPLSSNDRSKQKLRN VDRLNEQSQMDLDTIYTFEYPKAGGYFFSAPHTFSKIDHIIIGQKGLNRYKRIIEI IPCVLSIDHGLKLVENNNQGRNPITPDKLNALNLNDNI VKSEIKKEINSPLEPENENG		

Query Match	9.4%	Score 814.6	DB 3	Length 5133
Best Local Similarity	66.0%	Pred. No. 2.2e-10		
Matches 1129	Conservative	0	Mismatches 654	Indels 32
			Gaps 9	
QY	1	TGTATGAAGCCAAAGTCACATTAAATACCAAAACAGGAAAGATATA--CAAAAAGAA	58	
Db	2756	TCTAGGAAGCCACAATTACTTTATACCTAAACCAAGACACAAGAAAAGA	2815	
QY	59	ACTATGACGACGTACCACTGATGAATATACATGCAAAATCCCAAAATCTAGCTA	118	
Db	2816	ACTTCAGACCAATTTCCCTTATGAATATCGATGCAAAATCTCAACAAATTCCTGGCA	2875	
QY	119	ACCAATCCACAGATATCAAGAAGTAATCCACCATGTCAGTGGGCTTATPACAG	178	
Db	2876	ACCGAATCCAAAGACACATCAAAACAATCATCCACATGACCAAGTAGGCTTACCCAG	2935	
QY	179	GGGTGCA-GGATAGGTTAACTATACACAAGTAATATGATATCATCATCAATAACAGA	237	
Db	2936	GCAATGCGAGGATGTTATATATAGGAAACCAATCAACGTATCATTTATATTAACAAC	2995	
QY	238	TTAAAAACAATAATCATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT	297	
Db	2996	TGAAAGAACAAACCAACATGATCATTTCTTATAGATGCTGAAAGACATTTGACAAATTC	3055	
QY	298	AACATTTCTTATATATTAACCTTCAGCAAAATCGACATGAAAGACATACCTTATG	357	
Db	3056	AACACCCCTTATATATTAAGTCTCTGAAACAAATGGAATCAAGCCCATCACTGAACA	3115	
QY	358	TAAATAAGCATATATATAGACGACCCACAGCAAAACATTAATCTGATGGGAAAGTTGA	417	
Db	3116	TAGTAAAGCATATATACGCAAAACAGTGTCAATTAATTAATATGAGAGAAATCGC	3175	
QY	418	AAACATGTCCTGGAACCTGGAACAGACAGAGATG-CTACTTTCACCACTTATATCA	476	
Db	3116	AAGCAATCCACTAAATATCAAGGACATGACAAAGGCTGCCACTCTCTCTCACTTATTA	3235	
QY	477	ACATATGATGGAAGTTTATAGCCAGAGCAATCAACAAGAGAAAGAAATCAAGGCCACC	536	
Db	3236	ATATATGTTCTTGAAGTTCTACCCAGAGCAATCAAGCAACAAAGAGGTCAAGGGGATAC	3295	
QY	537	AAATCAATAAGAGGAAGTCAAACTGTCCCTGTTCACTGATGATATGATTTGATATCCAG	596	
Db	3296	AGATCGAAAGAAAGAAAGTCAAAATATCACTGTCTGACAGATGATATGATATGATTTTAA	3355	
QY	597	AAAAACCTTAAGAATCATCCAGAAAGCTCTCTAGAACATGATCATTAATTCAGTAAAGTTT	656	
Db	3356	GATATCCAAAGTTCACACAGAGAACTATTAAGCTGATTAACAATTCAGCAAAGTG	3415	
QY	657	CAGATACAAATCTAATATGACAAATCAATGATGACCTGCTATACCAACAGTACCAAG	716	
Db	3416	CTGGGTATTAATTAATCAATCAATTAATCAGTAGCTTCTCTTACACAAAGAGAAACAG	3475	
QY	717	CTGAGATCAAAATCAAGAACTCAAAACCTTTTCAATAGCTGTAA-----AAATATCT	770	
Db	3476	CGAGAAAGAAATTAAGGAAACGACACCTTCTATATATAGACCAACAAATTAATATTAAGTAC	3535	

Oy	771	TAAAGAAATTTCTTCCCAAGGAGGGTGAAGAGCCTCTAACAGAGAAAACCTACAAAACACAG-	829
Db	3536	TGGGTGTGACTTTATACAGACAGATGAAGTAAGTCTGTCAATAGAACTACAGACACTGCA	3595
Oy	830	---CTGACATCATAGATGATGACAAACAAGTGGAAAACATCCCATGCTCATGATGGGTA	886
Db	3596	AGAAGAAATTTGAAGAMAGACCTCGAAGATGGAAAGATCTCCCATGCTCATGATTTGGCA	3655
Oy	887	GAATCAATATTTGTGAANAATGACCAATTTGGCCAAAAGCAATCTACAGATTCGAATGCAATTC	946
Db	3656	GGATTTAATATGTAAATAAATGGCCATTTTACAAAAGCAATCTACAGATTCGAATGCAATTC	3715
Oy	947	CCACCAAAATATTCATCATCATATTTCTTCACAGAACTAGAAAACCAATCTAAAATTCATA	1006
Db	3716	CCATCAAAATATTCATCAATTCATTTCTTCAAAAGTTAG-ACAGAACATTTGGCAATTCATC	3774
Oy	1007	TGGAACAACAACCAAAAAAAAAAAAAAAAACCCCGATATGCCAAGCAAGACTTTCGAAA	1066
Db	3775	TGGATAATAC-----AAAAAACCGAGATGTAAAGCTATCTCTCAACATTA	3820
Oy	1067	AGAACAAATCTGSAGGATGACATTAACCATCTTGAAACTATPACTACAGAGCTATATCA	1126
Db	3821	AAAGGACTTCAGGGGGAAATCACTATCCCTGAATCAGACAGTATTTACAGAGCAATAGTGA	3880
Oy	1127	CCAAAACATCATGGCATGACATTAATAAATAAGGCATATAGCCAAATGAAAAAGAGAGA	1186
Db	3881	TAAAAACTGCCGTGTATTTGTATACAGACAGACAGACATAGACCAATGGAATGAATTTGAG	3940
Oy	1187	ATCCGAATAATTAAGCCAAATATTTATACCAACTGATTTTGGACAAACCAACAAAACA	1246
Db	3941	ACCCGAATAATTAAGCCCAACCTATATGTCTGATATTTTGGACAAAGAACCAAAACA	4000
Oy	1247	TAAAGTGGGAAAAAGACATTTCTAGTTAACAATGTGTGTGATTAATTTGGCAGGCCAT	1306
Db	4001	TCCATGTGAAAAAAGATATGACATTTTACGAATATGTGTGTGTCACTGAGAGTCAAT	4066
Oy	1307	GTGGAAGATGAATGATGATCCCTTGTCTCTCATTTAATACAAAATTGATACAGATGG	1366
Db	4061	GTAGAGAGATCAGATTCATCATAGCT-TATCACCTCTGTACAAAGCTTAAGTCCAAAGTG	4119
Oy	1367	ATCAAGACTTAATCTGAGACCTTAAACCAATTAATAATTTCTAGAGATTAACATCAGAAA	1426
Db	4120	ATCAGAGACTTCACATCAAAACAGACACACTCAATTAATAGAGAAAACCTAGGAGAG	4179
Oy	1427	ATGCTTCTAGACATTCACCTATAGGCAAGACCTCATGGCCAGAACCCAAAAGTAATGCA	1486
Db	4180	CATCTGGAACACATGGGCACTGGAAAAAATTTCTTAACAAAACCAATGGCTTAGCGCT	4239
Oy	1487	ACAAAAACAAAAATTAATATAGATAGACTTAATTAACCTAAAAAGTTTGGCGACGAAA	1546
Db	4240	CTAAGATTAAGATATGACAAATGGGATCTCAATAAAATGCAAAAGTTCTGTAGGCAAG	4299
Oy	1547	ACAATCATTTACAGACAGCAACAGAACCCACCGAGTACAGAAAAATCTTCACAAACTTA	1606
Db	4300	GACACTGTGTGTTAGACAAAAACGGAACCAACAGACTGGAAAAAGATCTTTACCAATCT	4355
Oy	1607	GCATCTGACTAAGGACTTAATATCCGGAATCCAAGAACTTCAACCAATCAGCAAGAG	1666
Db	4360	ACAACAGATTAAGGCTTATATCCAAAATTAACAAAGAACTCAAGAAATTAGACCGCAGG	4419
Oy	1667	AAAGCAACATCCCATGAAGAAGTGGGCTAAGGACATGAATATGACATTTCTCAAAATA	1726
Db	4420	GAAACAAATTAACCTTAATTAATAATGGGTTGACAGCTTAACCAAGAAATTCACAGCTAG	4479
Oy	1727	GATATTAACAATGGCCCAACAAACA--GGAATAAATGCTTAACATCACTAATGTTAGGAA	1784
Db	4480	GAATGCCGAATGGCTGAGAAACACTTAAGAAATGTTCAACATCTTTAGTCATAAGGAA	4539
Oy	1785	ATGTAAATCAACACTGTATATGATGATACCACTTAATCTCTGCAAGAAATGTCATATTTTA	1844
Db	4540	ATGCAAATCAAAACAACCTGTAGATTTCACTGCACACAGTGTAGATGGCTTAAGATCAAA	4599

ORIGIN

Query Match 9 3%; Score 801.8; DB 3; Length 3654;
 Best Local Similarity 65.6%; Pred. No. 1.5e-167;
 Matches 1321; Conservative 0; Mismatches 662; Indels 32; Gaps 9;

1 TGTATGAAGCCCATGCTCACTTTAATACCAAAACGAGAAAGATATAT--CAAAAAGAAA 58
 Db TCTATGAAGCCCATGCTCACTTTAATACCAAAACGAGAAAGATATAT 1060
 QY 59 ACTATGACCACTGACATGATGATATATATGACAGAAATCCCAAAATATCTAGCTA 118
 Db 1061 ACTTACAGACCAATTTCCCTTATGAAATATGACGCAAAAATATCTCAATTAATCTGCGCA 1120
 QY 119 ACCCAATCCACAGATATGCAAGAAATATCCACATGTCAGAGGGTTCATCCAG 178
 Db 1121 ACCGAATTCAGAGACATCAAAACCAATCATCCATGATATCAAGTAGGCTTCATCCAG 1180
 QY 179 GGGTGA--GATAGGTTAATATACATCAAGTCAATTAATGATATACATCAATTAAGAA 237
 Db 1181 GCATGACAGGATGCTTATATATACGAAAACATCAAGTATCCATTAATTAATTAACAAAC 1240
 QY 238 TTAATAACAAAATCAATGATATCTCAATGATGCTGAAAAAGCAATTTGACAAATCT 297
 Db 1241 TGAAGAACAAGAACACATGATATCTCAATGATGCTGAAAAAGCAATTTGACAAATCT 1300
 QY 298 AACATTTCTTATGATTAATACCTTCAGCAAAATGACATAGAAAGACATACCTTAATG 357
 Db 1301 AACACCCCTTCATGATTAATAGTCTGAAAAATGAAATCAAGGCCATACCTTAATG 1360
 QY 358 TAAATTAAGCCATATATGACGAGCCACAGCAAAACATTAATGAAATGGGAAAAAGTTGA 417
 Db 1361 TAGTAAAGCCATATATACAGCAACAGTTGCTAAACATTAATTAATGAGAAACCTTG 1420
 QY 418 AAAACATGTCCTTGAGAACTGGAACAGACAGAGANG--CTACTTTACCACTTCTATTC 476
 Db 1421 AAGCAATCCCATTAATATGAGGACATGACAAAGCTGCCCATCTCTCTACTTATTTCA 1480
 QY 477 ACATGATGAGAGTTTATGACGAGCAATGACAGAAAGAAAGAAATCAAGGGACCC 536
 Db 1481 ATATGTTCTTAAAGTTCTAGCCAGAGCAATGACACAAAGAAAGATTAAGGGATAC 1540
 QY 537 AAATCAATTAAGAGAGTCAAACTGCTGTTCTCATGATGATATGATGTTATACCTAG 596
 Db 1541 AGATGGAAGAAAGAGTCAAAATATCTATTTGCAAGTACATGATGATATTTAA 1600
 QY 597 AAAACCTTAAGACATCATCAAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTT 656
 Db 1601 GTGATCCCAAAAGTTCCACAGAGAACTTAAGGCTGATGAACAACTTCACAGCAAAAGTGG 1660
 QY 657 CAGGATACAACTTAATGTAACAATATGATGACAGTATACCAACAGTACAGCAAG 716
 Db 1661 CTGGGATTAATTAATCTCAATTAATATGATGCTCTCTCTATTAACAAAGAGAAACAG 1720
 QY 717 CTGAGATCAAAATCAAGAACTCAAACTTTTACAATAGCTGTAA-----AAAATACT 770
 Db 1721 CCGAAGAAAGAAATTAAGGAAAGCAACCTTCAATTAATAGCCAAATTAATTAAGTAC 1780
 QY 771 TAAAGATATTTTACCAAGAGGTGAAGAACTCTTACAAAGAAAGAAATTAACAAACAG- 829
 Db 1781 TCGGTGTGCTTAATCAAGAGAAAGTAAAGATCTGTAATTAAGAACTTCACAAACAGCTGA 1840
 QY 830 ---CTGACATCATAGTGAACAACAAGTGAAGAAACATCCGATGCTCATGAGGGTGA 886
 Db 1841 GGAAGAAATTAAGAGAGCTCTCAAGAGATGAGAAAGTCTCCATGCTCATGAGATTGGCA 1900
 QY 887 GAATCAATATTTGAAATGAACCATATTTGCAAAAGCAATCTCAAGTTCAATGCAATTC 946
 Db 1901 GGATTAATATGTAAGAAATGGCAATTTTACAAAGAAATCTCAAGTTCAATGCAATTC 1960
 QY 947 CCAACCAAAATATCATCATATTTTCAAGAACTAGAAAAAAACAATTTCAAAATTCATA 1006
 Db 1961 CCATCAAAATATCAATCAATCTTCTCAAGAGATTAG--ACAGAAACAATTTGCAAAATTCATC 2019

QY 1007 TGAACCAACCAACCAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGCTTAGCAAAA 1066
 Db 2020 TGAATTAAC-----AAAAAAAAAGAGATAGCTTAAGCTATCTTCAACATA 2065
 QY 1067 AGAACAAATCTGAGAGCATCACTTATCCATCTTCAAACTATCTACAGAGCTATATCA 1126
 Db 2066 AAAGACCTTCAGGGGAAATCACTATCTCTGAACTCAACAGATATACAGACCAATATGTA 2125
 QY 1127 CCAAAACATGAGCACTGACATTAATATGAGCAATAGCAACCAATGAAGAAAGAGGA 1166
 Db 2126 TAAAACTGATGTAATGTTGTAACAGACAGACATATGACCAATGAAATGAATTGAG 2185
 QY 1187 ATCCGAATTAATTAAGCCAAATTAATTAATGACCAACTGATTTTGAACAAGCAAAACA 1246
 Db 2186 ACCCAAAATGAACCAACACACTTATGTCATTTTGAACAAGAGCCAAACCA 2245
 QY 1247 TAAAGTGGGAAAAGACATCTTATTAACAAATGCTGAGATTATGCAAGCCACAT 1306
 Db 2246 TCCAAATGAAAAAAGATAGCATTTTACGCAATGCTGCTGTTCAACTGAGAGGCACAT 2305
 QY 1307 GTGAAAGATGAAATGATGATCCCTTGTCTCATCTTAATTAACAAATGATTAAGATG 1366
 Db 2306 GTAGAAAGATGAGATGATCATCTCT--TATCACCTGACAAACCTTAAGTTCAGTGG 2364
 QY 1367 ATCAAAAGCTTAATCTGAGACCTTAACCAATAAATTTAGAAATTAATCAATCAGAAA 1426
 Db 2365 ATCAAGAGACCTCAACATCAACCAACACACACTCAATTAATGAAGAAATCTAGGAG 2424
 QY 1427 ATGCTTCTAGACATTCATCTAGGCAAGACTTATGAGCCAGAACCCAAAGTAATGCA 1486
 Db 2425 CATCTGGAACACATGAGGCACTGAAAAAATTTCTGGAACAAACCAATAGGCTTATCT 2484
 QY 1487 AAAAAAACAATAATATATGATAGGCTTAATTAATTAATTAAGCTTTTGGCAGCAAAA 1546
 Db 2485 CTAGATCAAGATGACAAATGAGATCTCATTAATCTGCAAGCTTCTGTAAGGCAAG 2544
 QY 1547 ACATCATTAAGAGAGCAACCAAGACACCAAGAGTGAAGAAAAATCTTCAAACTTA 1606
 Db 2545 GACATGCTGTTAGAGCAAAACGAGCAACCAAGATGGAAGAAATCTTTACCAATCT 2604
 QY 1607 GCATCTGACTAAGACTAATATTCGGATTCACAGAACTCAACAAATCAAGCAAG 1666
 Db 2605 ACAACAGATAGAGGCTTATTTTCAAAATATTAAGAACTCAAGAAATGAGCCGAG 2664
 QY 1667 AAAGCAAAATCCCAATGAAGAGGAGTGAAGCAATGAATGAATTCACAAAGAA 1726
 Db 2665 GAAACAAATPACCTTAATTAATAATGAGGCTGAGCTTAACAAAGAAATTCACAGCTAG 2724
 QY 1727 GATATCAATAGGCAACAAACA--GAAAAAATGCTTAACATCACTAATGATTAAGGAA 1784
 Db 2725 GATGCGCAATGCTGAGAAACACCTTAAGAAATGTAACATCTTTGATTAAGGAA 2784
 QY 1785 ATGTAATCAACATGTAATGAGATACCACTTACTCTGCAAGATGCTAATTTAA 1844
 Db 2785 ATGCAATCAAAACAACCTGAGATTTCACCTTCAACAGCTGCTTAAGTCAAA 2844
 QY 1845 AAATCAAAATTAATGATGTTGAGGCTGAGTGAATTAAGAAACCTTTTACACTGC 1904
 Db 2845 AACTGAGTGAACGAGATGCTGAGAGATGAGAAAGAGAAACCTCTCTCAATGT 2904
 QY 1905 TGTGGAATGTAACTTGCAGCAACCTATGAAAAACAGTGTGAAAAATTTCTTAAGAA 1964
 Db 2905 TGTGGAATGTGAAGATGTTAAACATCTTGAATATGATGTGAGAGTTCTCTGAGAAA 2964
 QY 1965 CTAAAGATGATGACCATTTATTCAGCAATCC 1999
 Db 2965 TTGACATTTGAACCTGCTGAGATCAGCTATACC 2999

RESULT 5
 AY539883
 LOCUS

AY539883

4071 bp

mRNA

linear

HTC 21-MAR-2004

DEFINITION Rattus norvegicus LRG00132 mRNA, complete cds.
 ACCESSION AY539883
 VERSION AY539883.1 GI:45478103
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 1 (bases 1 to 4071)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B. Liver regeneration after PH

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 4071)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B. Direct Submission
 Submitted (03-FEB-2004) Henan Bioengineering Key Lab, Henan Normal University, NO.148 Jianshe Road, Xinxiang City, Henan 453002, P.R.China

FEATURES
 source
 CDS
 1..4071
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..4071
 /note="liver regeneration-related"
 /codon_start=1
 /product="LRG00132"
 /protein_id="AA566223.1"
 /db_xref="GI:45478104"
 /translation="MPTYTKLNNALLNDLVKBEIKKIKDFLENEEGTTPNLM DTKAVLRGFEISACRQKQERAYVSLTAHLKALEKQMPRRSRREIILKRAE FNOIEAVIKGLPTKRSRPGDSAEYPTIEDILPILSKLTHKIETDALSPTIE ATITLIPKPKDITKKNFRPISLMDAIIKILKILNRIQEHKTIYDQGFIPG MOGFNIRKTIIVIHINKLEQNHMIISDAKAPDKIQHPMKVLERIGIQGPY NIEIKIRAEINOVETRTIERINRTKSPFENKINKDKDLATRGRCEVOINKR NEKGIITDDEEIOKIRSYKMLSTPFNLOEMDFLRYOVSKLNOGIOLNPP ITPKEIAVIGKLPKTKSPGDSAEYPTIEDILPILSKLTHKIETDALSPTIE BATITLIPKPKDITKKNFRPISLMDAIIKILKILNRIQEHKTIYDQGFIP GMOGFNIRKTIIVIHINKLEQNHMIISDAKAPDKIQHPMKVLERIGIQGPY LNTVKAIVSKPVANIKLNGEKRLAIPKSGTQSCPLFNIIVLELRAIROQES IKGIQIKKEVKISLPADMIIVILSDPKSREBLKINNFSVAGKIKNSNSVAF YTEKQAEKEIRETTPPIDNNIKYLGVLTKOVOLYKQDLYPKEIKELIEEDLRK DLPSWIRGINMFMALIPKAIYRENAIPKIPLOPKELDPTICRPINNKPRJAK AILNNKTSGGITIPELKOYRAIVIKTAMTYRDRIOIDQWRIEDPENAPHYGHLI FDKGAKTIQWKDSIFSKWCFWNRATCRWQIDPSLPTCKLSKRIKDLHRTPL KILBEKLGKLEHMGKCNPLNTPMAVLRSLIDKDLIKQSFCARQIVATRKQ PDMEKIFTPNPTDRLGILSKIKYELKLCARFETNNPIKMGSELANEFTAKECMAEK HLKKGASLVIREMOIKTLRPHLTPRLAKIRNSGDRCMRGSGEGLTHGCMARQ LVPEFWKSWVRFLRKDILBPDPAIPLGIVPKDASTYRDCSTMEIALPLIARK KNEPRCPSTREMIQKMYIITYMEYSALKRNEPKYEGKLEGNITLSLTSQKDI YGEGSFWLTPVDOSVHGRKIMACSTFISNALPLKTIPTINPMGNISVIRKAGD LRDMDVLQIIQVMDVNWLOEGDMQMIIGVCLAPLKTATPTINPMGNISVIRKAGD GLIGPKENTIKKMFQSHNLKEDYEKKPQSRDEAF"

ORIGIN
 Query Match 9.3%; Score 801.8; DB 3; Length 4071;
 Best Local Similarity 65.6%; Pred. No. 1,56-167;
 Matches 1321; Conservative 0; Mismatches 662; Indels 32; Gaps 9;

1 TGTATGAGCGCAATGCTATTATACCAAAACAGAGAAAGATATAT--CAAAAAGAAA 58
 1346 TCTACAGAGCCACATTAATCTTATACCTTAACACACAAAGACACACAAAGAAAGAA 1405
 59 ACTATGAGCCAGTACCATGATGATATATACATGACGAATCCCCAACAAATATAGCTA 118
 1406 ACTTCAGACCAATATTCCTTATGATATGACGCAAAATAATCTCAATAAATTTCTGCA 1465
 119 ACCCAATCCACAGCATATCAAGAAATATATCCACCATTTGTCAAGTGGTTTCAATACAG 178

1466 ACCGAATTCAGAGCATATCAAAACATCATCAACCATGATCAAGTAGGCTTCAATCCAG 1525
 179 GGGTGA-GGATAGGTTAATACATACAAAGTCAATTAATGTATCATCATATAACAGAA 237
 1526 GCATGACGGATGGTTTAATATATGCGAAACATCAACGATGATCATATATTAACAAC 1585
 238 TTAATAACAAATACATATATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT 297
 1586 TGAAGAACAAGAACCATATGATCATTTCAATAGATGCTGAAAAAGCATTTGACAAATTC 1645
 298 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGAGCATCTTAATG 357
 1646 AACACCTTCATGATTAATAAGTCTGGAAGAAATAGAAATTCAGGCCATCACTTAACA 1705
 358 TAAATAAGCCATATATATGACGACCCACAGCAAACTTATATGTAATGGGAAAGTTGA 417
 1706 TAGTAAAGCCATATATACAGCAAAACGATGCTAATTAATTAATGAGAGAACTTG 1765
 418 AACATTTGCTTGAAGATGGAACAGCAAGAGATG-CTACTTTCACCACTTCTATTGA 476
 1766 AAGCAATCCCATTAATATGAGGACTAGACAGGCTGCCACATCTCTCCCTACTTATTCA 1825
 477 ACATAGTAGTGAAGTTTATGACAGGATATGACAGCAAGAGAAAGAAATCAAGGCACC 536
 1826 ATATAGTTCTTGAAGTTCTAGCCAGACATCAAGACAAAGAGATCAAAAGGATAC 1885
 537 AATATCAATAAGAGAGAGTCAAACTGCTCTGTTCACTGATGATATGATTTATACCTAG 596
 1886 AGATCGAAAGAGAGAGTCAAAATATCATATTTGAGATGATATATATATTTAA 1945
 597 AAAACCTTAAGACTCATCAAGAAAGCTCTGAACTGATATCATTAATTCAGTAAAGTTT 656
 1946 GTATATCCCAAAAGTTCCACAGAGAACTAATAAGCTGATTAACAATTCAGCAAGTGG 2005
 657 CAGATTCAGAACTAAATGACAAATCACTAGACACTGCTATACCAACAGTACCAAG 716
 2006 CTGGATATTAATAATTAATCAATTAATCAATGATGCTCTCTTATACAAAGAGAAACAG 2065
 717 CTGAGATCAATCAAGAACTCAAACTTTTCAATAGCTGTAA-----AAATACT 770
 2066 CCGAGAAAGAAATTAAGGAAACGACACCTTCTATATAGACCCAAATATATTAAGTACC 2125
 771 TAAGAATATTTTAACTCCAGAGAGTGAAGACCTTCAAGAGAAACCTACAAACACAG- 829
 2126 TCGGTGACTTAAACCAAGCAAGTAAAGATGTGTCATTAAGAACTTCAAGACACTGA 2185
 830 ----CTGACATCATAGATGACAAACAGTGAACACATCCCATGCTCATGTGGTGA 886
 2186 GGAAGAAATTAAGAGAGACCTCAGAGAGATGGAAGATCTCCCATGCTCATGTGGTGA 2245
 887 GAATCAATATTTGAAAAATGACCATATATGCGAAAGCAATCTCAAGTTCAATCAATTC 946
 2246 GGAATATATGTAATAATGCGATTTTACCAAAAGCAATCTCAAGTTCAATCAATTC 2305
 947 CCAACAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACAATTTCAAAATTCATA 1006
 2306 CCAATCAAAATATCAATCAATCTTCAAGAGTTAG-ACAGAACAATTTCAAAATTCATC 2364
 1007 TGAACACACACCAAAAAAAAAAAAAAAAACCCGCATAGCCCAAGCAAGCTTAGCAAAA 1066
 2365 TGAATATAC-----AAAAACCCAGAGATGCTAAAGCTATCTTCAACATATA 2410
 1067 AGAACAATCTGAGAGCATCATTAATCCATCTTCAACATATATCTAACAAGCTTAATCA 1126
 2411 AAGAGCTTAGGGGAAATCATATCCCTGAACCTCAAGCAGTATTCAGAGCAATATGTA 2470
 1127 CAAAAATCATGACCTGACATTAACATTAAGCAATGACCAATGAAAAAGAGAGAGA 1186
 2471 TAAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2530
 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCAATGATTTTGAACAAAGCAAAACAA 1246

Db 2531 ACCGAGAATGAGACCCACACACCTATGCTGATTTTGAACAAGGACCAAAACCA 2590
 Qy 1247 TAAAGTGGGGAAGAGACATTTAGTTAACTATGCTGAGATTTTGGCAAGCCACT 1306
 Db 2591 TCCATGGAAGAAAAAGATGACATTTTCAGCAAAATGGTGGTTCACATGGAGGCAACAT 2650
 Qy 1307 GTGAGAGATGAATGAACCTGATCCCTGCTGCTCACTTAAATACAAAATTGATCAAGATGG 1366
 Db 2651 GTAGAGATGACAGATCAATCATCTCT-TATCACTCTGTAACAAAGTTAAAGTTCAGTGG 2709
 Qy 1367 ATCAAGAATTAAATCTGAGACCTAAACCATTAATAATTTCTAGAGATTAACATGAGAAA 1426
 Db 2710 ATCAAGAGCTTCCATCATCAACACAGACACTCAAACTAATAGAGAAAACTAGGGAAG 2769
 Qy 1427 ATGCTTGAACATTCATCTAGGCAAGAACTTCAATGGCCCAAGAACCCAAAGTAAATGCA 1486
 Db 2770 CATCTGGAACACATGGGCACTGGGAAAAAATTTCTCGAACAAAACCAATGGCTTAATGCT 2829
 Qy 1487 ACAAAAACAAAATTAATAGATAGAGACTTAATTAATACTAAAAAGTTTGGCGAGCAAAA 1546
 Db 2830 CTAAAGATTAAGATGACAAATAGGATCTCATTAATCTCAAAAGCTTCTGTAAAGCAAG 2889
 Qy 1547 ACAATCATTAAGACAGACAAACAGACACCACGAGTGAAGAAAAATCTTCACAAACTAA 1606
 Db 2890 GACACTGTGTTAGGACAAAACGGAACCAACAGATTTGGGAAAAAGATCTTTACATCTCT 2949
 Qy 1607 GCATCTGACTAAGACTAATATCCGGAATCCCAAGAACTCAACAAATCAGCAAGAG 1666
 Db 2950 ACAACAGATGAGGCGCTTAATTTCCAAAATATCAAGAACTCAAGAAATTAAGCCGACAG 3009
 Qy 1667 AAGAGAAACAATCCCATGAAAGAGTGGGCTAAGACATGAATGACAAATTCCTCAAAAAGA 1726
 Db 3010 GAACAAATTAACCTTAATTAATAATGGGTTAGAGCTTAAACAAAGATTCACAGCTAAG 3069
 Qy 1727 GATATACAAATGCGCAACAACA--GGAATAATGCTTAACATCACTAATGATTAGGAA 1784
 Db 3070 GAATGCCAATGCGCGAAGAAACACCTTAAGAAATGTTGACATCTTATGATTAAGGAA 3129
 Qy 1785 ATGTAATCAACACTGTATGTGATACCACTTAATCTCTGCAAGATTTGCTCAATTTAA 1844
 Db 3130 ATGCAAAATCAAAACAACCTGAGATTTCACTCAACAGATGGATTTGCTAAGATCAAA 3189
 Qy 1845 AAATCTAAATAATATAGATGTTGGTGGTCTGTGGTGAATAAGAACTTTTCACTGAC 1904
 Db 3190 AACTCAGGTGACAGAGATGCTGGGAGATGTGAGAAAGAGAACATCTCTCCATTTGT 3249
 Qy 1905 TGGTGGAGATGTAACCTTGCGCAACCACTATGAGAAAAAGTGTGAATTTCTTAAGAA 1964
 Db 3250 TGGTGGAGATGTCAGACTGGTAAACCATTTCTGGAATTCAGTCTGAGAGTTCTCAGAAA 3309
 Qy 1965 CTAAAGATGATGACATTTGATTCAGCAATCCC 1999
 Db 3310 TTGGACATTTGAATGCTGAGAGATTCAGCTATATAC 3344

RESULT 6
 AY539886
 LOCUS AY539886 4356 bp mRNA linear HTC 21-MAR-2004
 DEFINITION Rattus norvegicus LRRG00135 mRNA, complete cds.
 ACCESSION AY539886
 VERSION AY539886.1 GI:45478109
 KEYWORDS HTC
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 4356)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xing, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Liver regeneration after PH
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4356)
 AUTHORS Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xing, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2004) Henan Bioengineering Key Lab, Henan Normal
 University, NO.148 Jianshe Road, Xinxiang City, Henan 453002,
 P.R.China

FEATURES
 source
 1..4356
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..4356
 /note="Liver regeneration-related"
 /codon_start=1
 /product="LRRG00135"
 /protein_id="AA566226.1"
 /db_xref="GI:45478110"
 /translation="MNIGSNHYSLISLNLINGNSPIKRRHLTNMIGNEDPACDQ
 EPHLEDRHYLRVKGWKTTRQANQOKQAGAVLISNKLINPOLKVIKKQEGHFI
 KKHIDELSLNTIYAPTPATPYTKETLKLKTHIAHTIITVDPNTPLSMSRSMK
 OKLNSDVLRLEWVSQMDLTDIYRTPEKAGVTFPSAPHGTSKIDHTIGQKGLK
 YRKLEIPCVLSDHGLKLVFNNNKGRMPTWTMLNMLNDNLVKEIKKEINFLB
 FNEBEDTTPNLMIDTMKAVLRGKLIALSACRKOBRAYVSLTAHLALBQKANTPR
 RSRROEIKLRAEINOVETKRTIERINRTKMPFEKINKIDKPLARLTRHRECOIN
 KIRNKKDITDSEIOLKIRSYKNISTPEMLQEMDNFLDRYOUSKINQOINL
 NNPIPKIEIVIKGLPTKSPGPGPGEAEYQPIBELIPIISLKHPIKHTIDALRN
 SEYEATITLIPKSHDITKENFRISLMDLIDAKILNKILNRIOEHKITHDOQG
 PLTQGWENIRKMTNVPFHYINKLEQNMILSLDAEKAFDIIQHPMIKVLBIQIO
 GYELINIVAIYSKPVANKLNGEKQIALPLSGRQGGPLSPYELNFIKVLAEIAIRK
 KEVKGIOIGDEVKISLPADMDIAYLSDPSGTRRELKLNFIKVGKYLKNSKSV
 AFLYTKROAEKEIRETTPIIIDPNVIITLKVYDITKQYDLYNKGKILREDEDR
 RKMDPCSGIRGINVKAAILPKAYRNALPIKPIOPPELDRITCKFIWNKKPR
 IAKAILNNKRTSGGTTIPELKOYTAIYIKTAMWYRROIDQNRIDEEMNPTTG
 HLIFDGAKTIOWKDLSFSKCPWPNWATCRMOIDCLSPCTKLKSKWKIDHTIP
 DTLKLEKLGHRHMGSTGRKFLKTPMAVALRSRIDKMLIKLQSCAKADIVRT
 KQPTDWEKIFTNPTTREGLSIKYELKLDRETNTPIKMGSELKRETFACBWA
 AKELKKCSTSLIIREMOIKTTLRPHLPVPLAKIKNSGDSRCWRGCEBGTLLHCWM
 DCLRYKPRKSWRLRLKLDIELRDPAILGIYIPDASVYKRDSTMTIALFTI
 ANMKPRCPSTEWRTIQKMYIYTHEYSAIKNNEPMKPVGKMLEBNTIISLITQSO
 KOIHDLVOTCLMEHCIDQCMWGGSDHICQAMNATYDPTQGRNFKLCIPIHSITRYV
 LAKPQEGCAMNDQGRGDELLALASGVNRGCDPLQGGCQALAKEPRLALBSWP
 IRCPAGHPVMELEWGNPPPPPPHPSDPCGVTRVVEBGTFRFVGDLLILGWGVG
 SYAARQLPCCALTIY"
 ORIGIN
 Query Match 9.2%; Score 790.6; DB 3; Length 4356;
 Best Local Similarity 65.2%; Pred. No. 4,9e-165;
 Matches 1314; Conservative 0; Mismatches 669; Indels 32; Gaps 9;

Qy 1 TGTATGAAGCCATGCTACTTTAATACCAAAACCGAAAGATATA--CAAAAAGAA 58
 Db 1529 TCTACAGACCAACAATTACTCTTAATACCTAAATGACAAAGACCAACAAAGAAAGGA 1588
 Qy 59 ACTATGACAGATGACATGATGATATATCATGACGAAGATCCCAACAAATACTAGCTA 118
 Db 1589 ACTTGACACCAATTTCCCTTAATGATATGACGACAAATTAATCTCAATAAATTCGGCA 1648
 Qy 119 ACCCATCCAGACATATCAAGAAAGATATCCACATTTGCAAGTGGGTTTATACAG 178
 Db 1649 ACGAATTCAGAGCAGATCAAAACATCATCATCAATGACCAAGTAGGCTTCATCAGAG 1708
 Qy 179 GGGTGA--GATAGGTTAATCATACACAGTCAATTAATGTGATATCATCAATTAACAGAA 237
 Db 1709 GCATGACAGGATGGTTAATATACGAAACCATGACGATGTTTATATATTAACCAAT 1768
 Qy 238 TTAATAACAAAATACATGATCATCTCAATGATGCTGAAAAGACATTTGCAAAATCT 297
 Db 1769 TGAAGAAACAAACCATGATCATTTTCAATGAAGCTGAGAAAGACATTTGACATTAATC 1828
 Qy 298 AACATTTCTTATGATTAACCTTCAGCAAAATGACATGAGAAAGACATACCTTAATG 357

Db 1829 AACACCCCTTCATGATAAAGTCTGGAAAAGATTGGAATTCAAGGCCCACTGTAACA 1888
 Qy 358 TAATAAAGCCATATATGAGGAGCCACAGCAAACTTATATCTGAATGGGAAAAGTTGA 417
 Db 1889 TAGTAAAGCCATATATGAGGAGCCACAGTGTCTTAACTTAAATGAGAGAAATC 1948
 Qy 418 AACATGTCTCGAGAACTGGAAACAAGAGAGT-CTACTTCCACCTTATTTCA 476
 Db 1949 AAGCAATCCCACTTAAATCAAGGAGCTTGAACAAGGCTCCCACTCTCTCTTATTTCA 2008
 Qy 477 ACATAGTAGTGAAGTTTGAAGCCAGACAAATCAGACAGAGAAAGAAATCAAGGCAACC 536
 Db 2009 ATATAGTTCTGAAGTTCTAGCCAGAGCAATCAGAAAAGAGAGGTCAAGGGATAC 2068
 Qy 537 AAATCAATTAAGAGAGAGTCAAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 596
 Db 2069 AGATCGGAAAAGAGAGAGTCAAAATATCACTGTGTGAGATGATATGATATGATTTAA 2128
 Qy 597 AAAACCTTAAGACTCATCCAGAAAGCTCTAGAAAGTATCATTAATTCAGTAAGTTT 656
 Db 2129 GTGATCCCAAAAGTTCCACAGAGAACTTAAAGCTGATTAACAACTTCATCAAGTGG 2188
 Qy 657 CAGATCAAACTAAATGTAACAATTCAGTACGCTCTATACCAACAGTACCAAG 716
 Db 2189 CTGGGTATTAATTAATCAATCAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2248
 Qy 717 CTGGAATCAATCAAGAACTCAAACTTTTACATATGCTGTAA-----AAAATACT 770
 Db 2249 CCGAAGAAAGAAATTAAGGAAACGACCTTCATTAATAGACCCAAATTAATTAAGTACC 2308
 Qy 771 TAAGAATATCTTAAACCAAGAGGTGAAGAGCTCTCAAGAAAGAACTTACAAACAG- 829
 Db 2309 TCGGTGACTTTAAACAGAGAGTAAAGTCTGTACAAATGAACTTCAAGCACTGA 2368
 Qy 830 ---CTGACATCATGATGACACAAACAAGTGAACAATCCCTGCTCATGATGGGTA 886
 Db 2369 GGAAGAAATTAAGAGAGAGCTGAGAGATGAGAAAGATCTCCATGCTCATGATGGCA 2428
 Qy 887 GAATCAATTTGGAAGAAATGACATATTTGCCAAAAGCAATTCATCAATGCAATTC 946
 Db 2429 GGAATTAATTAAGTAAATGAGGCACTTTTACAAAAGCAATTCATCAATGCAATTC 2488
 Qy 947 CCAACCAAAATATCATCATCTTCTTCAAGAACTAGAAAACAAATTTCAAAATTTATA 1006
 Db 2489 CCAATCAAAATATCAATCAATCTTCAAGAGTTAG-ACGAAACAAATTTCAAAATTTATC 2547
 Qy 1007 TGAACCAACCAACCAAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCA 1066
 Db 2548 TGAATTAAC-----AAAAAACCCAGAGATAGCTAAAGCTATCTCAACATTA 2593
 Qy 1067 AGAACAATCTGAGGATCATTAATCCATCTTCAAACTATATCTACAGGCTATTAATCA 1126
 Db 2594 AAGGACTTCAGGGGAGATCATATCCCTGAACCTCAAGAGTATTTACAGAGCAATAGTGA 2653
 Qy 1127 CCAAAACATGATGACATGACATTAATCTAGGCACTAGACCAATGAGAAAGAGAGA 1186
 Db 2654 TAAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2713
 Qy 1187 ATCCAGAAATTAAGCCAAATTAATTAATAGCCAACTGATTTTGAACAAAGCAACAAACA 1246
 Db 2714 ACCCAAGAAATTAAGCCCAACCTATAGTCACTTGAATTTTGAACAAAGAGCCAAACCA 2773
 Qy 1247 TAAAGTGGGAAAGACATTTCTAGTTAAACAATGCTGAGATTAATTTGCAAGCCCAT 1306
 Db 2774 TCCAAATGGAAGAAAGATAGATTTTTCAGAAAGTGTCTGTTCACTGAGGCAACAT 2833
 Qy 1307 GTGGAAGATGAAGTGAATCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1366
 Db 2834 GTAGAGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2892
 Qy 1367 ATCAAAAGCTTAATCTGAGAGCTTAAGCAATTAAGCAATTAAGCAATTAAGCAATTA 1426
 Db 2893 ATCAAGAGCTTCAATCAATCAACAGACACTCAAACTAATTAAGAAAGAACTAGGAGAG 2952

Qy 1427 ATGCTTTAGACATTTCACTTAGCAAAAGACTTCAATGCCCAGAAAGCCAAAGTAAATGCA 1486
 Db 2953 CATAGGAGACATGAGGACCTGGAAAAATTTCTGTGAACAAACCAATAGGCTTATGCT 3012
 Qy 1487 ACAAACCAAAATTAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 3013 CTATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 3072
 Qy 1547 ACATCTTAGACAGAGCAACAGACCAACCAAGTGAAGAGAAATCTTCAACACTTA 1606
 Db 3073 GAGCTGTGTTAGGACAAAACGGCAACCAAGATTTGGGAAAAGATCTTTACCAATCT 3122
 Qy 1607 GCATCTGATTAAGGATTAATTCGGATTCACAGAACTCAACAAATCAGAGAG 1666
 Db 3133 ACAACAGATGAGGCTTATATCAAAATATACAAAGAACTCAAGAAAGTTAGACCGCAG 3192
 Qy 1667 AAGCAACCAATCCATGAAGAGTGGGCTTAAGGACATGAATGACAATCTCAAAAGAA 1726
 Db 3193 GAAACCAATTAACCTTATTAAGAAATGGGTTCAAGCTTAACAAAGAAATTCACAGCTGAG 3252
 Qy 1727 GATATCAAAATGCGCAACAAACA--GGAAAAATGCTTAACTATCAATGATTAAGGAA 1784
 Db 3253 GAATGCCGAATGCTGAGAAACCACTTAAGAAATGTTCAATCTTTATATCATTAAGGAA 3312
 Qy 1785 ATGTAATCAACACTGTAATGCGATACCACTTATCTCTCGAAGATGCTATTAATTTAA 1844
 Db 3313 ATGCAATCAAAACCAACCTGAGATTTCACTCAACAGTGGCATTTAGATCAAA 3372
 Qy 1845 AAATCTTAATTAATTAATGATGTTGGTGTCTGTGTGATTAAGAAACATTTTACATGC 1904
 Db 3373 AACTGAGTGAACGACATGCTGCGAGATGAGAGAAAGAGAAACATCTTCATTTGT 3432
 Qy 1905 TGGTGGAAATGTAACCTTGGCAACCACTATGGAACAAAGTGTGGAATTTCTTAAGAA 1964
 Db 3433 TGGTGGATTTGACAGCTGTAACCAATCTTGGAAACATCTGGAAGTTCTTCAAGAAA 3492
 Qy 1965 CTAAAGTAGATGACCATTTGATTCAGCAATCCC 1999
 Db 3493 TTGGACATTTGAACCTGCTGAGATCCAGCTATACC 3527

RESULT 7
 BC038425 2488 bp mRNA linear HTC 04-MAR-2003
 LOCUS
 DEFINITION Homo sapiens, clone IMAGE:4818574, mRNA.
 ACCESSION BC038425
 VERSION BC038425.1 GI:23512256
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2488)
 Strausberg, R.
 Direct Submission
 Submitted (01-OCT-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiroki
 Tohyuki and Piero Carninci (RIKEN)
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: ang@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

Kowis, C.R.; Sneed, A.J.; Martin, R.G.; Muzny, D.M.; Nanavati, A.N.; Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.linl.gov>
 Series: IRAP Plate: 71 Row: a Column: 12
 This clone has the following problem: retained intron.

FEATURES	Location/Qualifiers
source	1. . 2488

```

/organisms="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clome="IMAGE:818574"
/tissue_type="Brain, hippocampus"
/clome_ids="NIH MGC_95"
/lab_host="MDH10B"
note="Vector: pBluescript"

```

ORIGIN

Query Match	8.9%	Score 770;	DB 3;	Length 2488;
Best Local Similarity	83.7%;	Pred. No.1.8e-160;		
Matches 948;	Conservative 0;	Mismatches 165;	Indels 19;	Gaps 6;

Qy	872	GCTATGATGGGTGATTCATATATTTGGAAAAATGACATATTTGGCAAAAGCATCTCA	931
Db	1325	GCTATGGATAGTGTGAATTCATATTTTGGAAAATGACATCTCTTAAAAGCATCTCA	1388
Qy	932	AGTTCAATGCAATTCACAATAATATCATCATTTCTCCAGAACCTAGAAAAAACA	991
Db	1385	AATTCACAACAATTCCTCAATCAAAAATACCAACCATCATTTCTCCAGAAATTAG-AGAAAACA	1443
Qy	992	ATTCTAAAATTCAATTGGAAACAACAACCAAAAAAAAAAAAAAAAAAACCCTGATGGCAAG	1051
Db	1444	ATTATAAAATTCATATGGAAC-----CAAAAAAGACCCACATATCCAAG	1489
Qy	1052	CAAGACTTAGCAAAAAAGAACAAATCTGAGGATCACATTCACCTCTTCAACATATCT	1111
Db	1490	CAATATCTAGCAAAAAGAACAAATCTGAGGATCACATTTACTCTGATTTCAACATATGCT	1549
Qy	1112	ACAAAGCTATATACCAAAACATCATGSCACTGACATPAAACTNAGGCACATNAGCAAT	1177
Db	1550	ATTAAGGCATATGTCCAAAAACAGCATGGTACTGGTATPAAAAAAGAGGCAATNAGCAAT	1609
Qy	1172	GGAAGAAGAGAGAAATCCAGAAATPAAAGCCAAATPATTATATGSCAACATGATTTTGA	1231
Db	1610	GGACAGAAAAGAGAACTCAGAAATPAAACCAAAATCTTACAGCAATTGATCTTTGACA	1665
Qy	1232	AAGCAAAACAAAACATPAAAGTGGGAAAAGACATTTCTAGTTAAACAATGGTGCTGAGATT	1291
Db	1670	AAGCAAAACAAAACATPAAAGTGGGAAAAGACATCTTTTCAACAAATGGTGCTGGAGTA	1722
Qy	1292	ATTGGCAAGCCCATGTGTGAGAAATGAAACTGGATTCCTGTCTCTCACTTAATACAAA	1351
Db	1730	ATTGGCTTAGCCCATPACAGGAATAATGAAACTGGAT-CTTCATCTCTCACCTTAATACAAA	1788
Qy	1352	ATTGATPACAAAGATGATCCAAAGCTTAAATCTGAGACCTPAAACCATATAAAATTTAGAA	1411
Db	1789	ATCAACTCAAGATGATTTAGACATT-ATCTPAGACCTGAAACTPACAAAATTTCTAGAA	1841
Qy	1412	GATTAACATCAGAAAAATGCTTTAGACATTTCACTTAGGCAAAAGACTTCATGGCCAGAAC	1477
Db	1848	GATATCATTTGGCAAAACCCTTTACACATTTGGCTTAGGCAAGATTTCTAATACAAAGAAC	1907
Qy	1472	CCAAAAGTAAAAGCAACAAAACAAAATAATATAGATGGACTTAACTTAAACTPAAAAGC	1531
Db	1908	CTAAAAGCAAAATGCAATPAAACCAAAAGATPATT-GCTGGAGACTTAATTAACCTPAAAGAGC	1966
Qy	1532	TTTTTCGCGAGCAAAAACATCATTTAGCAGAGAGAAACAGACAACCAACCGAGTGAAGAAA	1591
Db	1967	TTTTCTCATAGGCAAAAAGAACATCTCAGCAGAGTAAAGACAAACCAACAGAGTGAAGAAA	2022
Qy	1592	ATCTTACAAACTPAAAGCATCTGACTPAGACTPATAATCCGGAATCCACAAAGAACTCAAA	1651

Db	2027	ACCTTCACAATCTATATGCTGACAAAGAGACATAATATCCGAATCTATTAAGAGCTCAAA	2086
Oy	1652	CAATTCAGCAAGAAAGAAACAAATCCCATGAAAGTGGGCTTAAGGACATGAATGAA	1711
Db	2087	CAAAATTGCGAAGAA - AAAAAAAAAAGATCCCATCAAAAAGTGGGCTTAAGGACATGAATGAA	2145
Oy	1712	CAATTCCAAAAAGAGATATACAATATGGCCCAACAAACAGAAAAAATAGCTTAACATCACT	1771
Db	2146	CAATTTCCCAAAAAGAGATATCAATATGGCCCAACAAACATATAAAAAATGCCAATCACT	2205
Oy	1772	AATGATTAAGGAAATGTAAATCAA CACTGTATGCGATACCACTTACTCTCGCAAGAT	1831
Db	2206	AATGATCAGGAAATGCAATCAAAACCAATGTGATACCACTTACTCTCGCAAGAT	2265
Oy	1832	GCGTATATATTTAAAAATCTAAAAATATAGATGTTGGTGGGTCTGTGTGTTAAAGAAC	1891
Db	2266	GGCCATATATCTTAAAAATCGAAAAATAGTAGATGTTGGCATGGAATGTGTGTATCGGGAAC	2325
Oy	1892	ACCTTTACATCGCTGGTGGGAATGTAAACCTTGCGCAACCATATGAGAAAA CAGTGTGAAA	1951
Db	2326	ACCTTCAAATGGCTGGTGGGAATGTAAACTATGATAGCCACTATAGAAAAACATGTGGAG	2385
Oy	1952	ATTCTTAAAGAACTAAAAAGTAGATTCGACATTTGATTCAGACAAATCCCATTA	2003
Db	2386	ATTCTTAAAGAACTAAAAAGTAGAAATCTACATTTGATTCAGACAAATCCCACTTA	2437

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION
HSMB04900	HSMB04900	2231 bp mRNA linear	HTC 03-AUG-2004	
		Homo sapiens mRNA; cDNA	DRFZp86C0964 (from clone	
		AL33587	DRFZp86C0964).	
		AL33587.1	GI:21734233	

REFERENCE	1 (bases 1 to 2231)
AUTHORS	Bahr, A., Lamber, J., Mewes, H.W., Well, B., Amlid, C., Osanger, A., Fodor, G., Han, M. and Wiemann, S.
CONSTRM	The German cDNA Consortium
TITLE	Direct Submission
JOURNAL	Submitted (13-JUL-2004) MTPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center, Heidelberg, Germany

FEATURES	Location/Qualifiers
source	1. .2231

	ORIGIN
Query Match	8.9%; Score 769.2; DB 3; Length 2231;
Best Local Similarity	73.5%; Pred. No. 2.6e-160;
Matches 1241; Conservative	0; Mismatches 263; Indels 184; Gaps 13


```

QY 864 CATCCATGCTCATGATGGGTAGATCAATATTGTAATAATGACCATATTGCCAAGC 923
Db 318 CCTCCCATGCTTATGATGGGTAGATCAATATTGTAATAATGACCATATTGCCAAGC 377
QY 924 AATCTCAAGTTCATGCAATTTCCCAACCAAAATTCATCATCTTCTTCAAGAACTGA 983
Db 378 AATCTCAAGTTCATGCAATTTCCCAACCAAAATTCATCATCTTCTTCAAGAACTGA -C 436
QY 984 AAAAAAATTTCTAAATTCATGTAACCAACCAAAAAAATTTCTTCAAGAACTGA 1043
Db 437 AAAAAAATTTCTAAATTCATGTAACCAACCAAAAAAATTTCTTCAAGAACTGA -C 481
QY 1044 AGCCAAAGCAAGCTTGCAAAAAAATTCGAGGATCATTAACCATTTCAA 1103
Db 482 -----TAGCAAAAAAATTCGAGGATCATTAACCATTTCAA 526
QY 1104 ACTATCTACAGGCTTATATTCACCAAAATTCATGATGATCAATTAATGACCAT 1163
Db 527 ACTATCTATAGGCTTATATTCACCAAAATTCATGATGATCAATTAATGACCAT 586
QY 1164 AGACCAATGAAAGAAAGAAATTCAGAAATTAAGCCAAATTAATTAAGCCAACTGAT 1223
Db 587 AGACCAATGAAAGAAATTCAGAAATTAAGCCAAATTAATTAAGCCAACTGAT 646
QY 1224 TTTTGAACAAGCAACAATTAAGTGGGAAAGCACTTTCTA----- 1269
Db 647 CTTTGAACAAGCAACAATTAAGTGGGAAAGCACTTTCTA----- 1269
QY 1270 GTTAAACAATGATGCTGATTAATTTGCAAGCCATGATGAAATGAAATGATGATCC 1329
Db 706 ACACCTTTTGTGCTGGGCTATTTGCTAGCCATGATGAAATGAAATGATGATCC -CC 764
QY 1330 TTGTCTCTCACTTAATTAATTAATTTGATCAAGATGATCAAGATTAATTTGATGACC 1389
Db 765 TCATCTCTCACTTTGATCAAAATTCATCAAGATGATTAAGGATTAATTAATTAAGACC 824
QY 1390 TAAACAATTAATTTCTAAGATTAATTCATCAAGAAATGCTTAAACATTTCTGATGACC 1449
Db 825 TGAACAATTAATTTCTAAGATTAATTCATCAAGAAATGCTTAAACATTTGATGACC 884
QY 1450 CAAGAATCTGAGGCAAGAAATTCATCAAGAAATGATCAAGAAATTAATTAATGATA 1509
Db 885 CAAGAATCTGAGGCAAGAAATTCATCAAGAAATGATCAAGAAATTAATTAATGATA 944
QY 1510 GGAATTAATTAATTAATTAATTTGCGAGCAAAAAATCAATCAATGACAGCAAGCAAG 1569
Db 945 GGGGCTTAATTAATTAATTAATTTGCGAGCAAAAAATCAATCAATGACAGCAAGCAAG 1004
QY 1570 ACAACCAAGGATGAGAAATTTCTCAAACTTAAGCATCTGATTAAGGATTAATTAATC 1629
Db 1005 AAAACCAAGGATGAGAAATTTCTCAAACTTAAGCATCTGATTAAGGATTAATTAATC 1064
QY 1630 CGGAATCTCAAGAAATTCAGAAATTCAGAAATGAAAGCAAAATCCCAAGAAAGA 1689
Db 1065 CAGAAATCTCAAGAAATTCAGAAATTCAGAAATGAAAGCAAAATCCCAAGAAAGA 1124
QY 1690 GTGGGCTAAGGATGATGATCAATTTCTCAAAAGAAATTAAGGATGATGATGATGATG 1749
Db 1125 GTGGGCTAAGGATGATGATCAATTTCTCAAAAGAAATTAAGGATGATGATGATGATG 1184
QY 1750 G--GAAAAATGCTTAATCACTAATGATTAAGGAAATGATTAATCAACATGATGATG 1807
Db 1185 TATGAATAATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244
QY 1808 ATACCACTTACTCTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1865
Db 1245 ATATCACTTACTCTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1304
QY 1866 TGGTGGTCTGCTGATTAAGAACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1925
Db 1305 TGGGCTGATGCTGCTGATTAAGAACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364
QY 1926 CAACCACTATGAAACAGTGTGAATTTCTTAAGAACTTAAGATGATGATGATGATGATG 1985

```

```

Db 1365 CAGCACTATGAAAC--TGTGAAGTTCTTTAAGAACTTAAGATTAAGTAACTACATTT 1422
QY 1986 GATTCAGCAATTCCTATTA----- 2003
Db 1423 GATTCAGCAATTCCTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1482
QY 2004 ----- 2003
Db 1483 AGATCTGCAACATATTTATGAGCAAAAAATTCAGAAATGCAAAATGCTGGAACCA 1542
QY 2004 -----AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2042
Db 1543 ACCCAATGCCCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1602
QY 2043 AACTCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2102
Db 1603 TATGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1662
QY 2103 ACCCTTAATTTAAGTGGGCTAATCTCAGAAATGCAAAATCAATCAATGATGATGATGAT 2162
Db 1663 ACTATTAATTTAAGTGGGCTAATCTCAGAAATGCAAAATCAATCAATGATGATGATGATG 1722
QY 2163 ACAAGTGGGCTAAGCTGAGGACAGAAAGCAT--AGATGATTAATTAATTAATTAATTAAT 2221
Db 1723 ATATGTTGAGCTTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1782
QY 2222 GACTTGAAGGGAAGATGAGAAAGAGGAGGAGGATTAAGATTAAGATTAAGATTAAGATTA 2281
Db 1783 GACTTGAAGGGAAGATGAGAAAGAGGAGGAGGATTAAGATTAAGATTAAGATTAAGATTA 1842
QY 2282 TACACTGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2341
Db 1843 TA--TACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG --- 1895
QY 2342 ATGGAAGCAAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2401
Db 1896 ATGGAAGCAAAATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1955
QY 2402 TGATTTAA 2409
Db 1956 AATTTTAA 1963

```

```

QY RESULT 9
AY383712
LOCUS 6966 bp mRNA linear HTC 15-OCT-2003
DEFINITION Rattus norvegicus LR8GT00057 mRNA, complete cds.
ACCESSION AY383712
VERSION AY383712.1 GI:37654329
KEYWORDS HTC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 6966)
Yang, K.J., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, D.Y.,
Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
Liver regeneration after PH
Unpublished
2 (bases 1 to 6966)
Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xing, X.K.,
Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
Direct Submission
Submitted (09-SEP-2003) Henan Bioengineering Key Lab, Henan Normal
University, No. 148 Jianshe Road, Xinxiang City, Henan 453002, P.R.
China
FEATURES
source
Location/Qualifiers
1..6966
/organism="Rattus norvegicus"

```

CDs

/mol type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..6966
/note="Liver regeneration-related"
/codon_start=1
/product="LRG00057"
/protein_id="AA096270.1"
/db_xref="GI:37654330"
/translation="MVLGVVASLAACPIWHPSPVVDITKGVLSGSEFT
QVAVFLGVPPAKPLGLFAPBPBAPBMSFVKVITTPYTPGSHSGNAHLDOL
AALRWQDNIAFGNPDVITFESAGAGSVSLMIALSCGNTTNSAAMQCLQK
TEALBELTVDKNTSMSTVIGVLPKTEELTEKSTNTVYIYAFKQWGMILP
TNISBSVPAIIEKLRGTHDPBPKKELLDMFSDFPRLPIAVMSRGLRDGAPDM
YEFQYRPSFVDHDPQYVQDGDDEIFSVFTPLFKEDGLEDIPDFSSVPSADLSR
FLQTLAVFLCEALCSFLQGECEQSDSEKQKQSKVGLCEDIPRPRGAMDE
HGQOQSKLMRLKYKASGLPWGRODCLBPNRLKEDMINSILNHPGWRGKAP
SERQSLNQKRLPAHTRSPBEKAKDLBPMCTEAPRGGTGLPGCCRCPELGST
PRANSLGTTEIAEQSNTRDMARGKRNPNRNDWPSSEPNPPTKTNMEYPTP
EKQDVSYSYLIMLTPDKDMWTLRETOENLNKOYAEARBESQSLKEPOENTIKOL
KELKEIETAIKEHEMTTLDINQKQGVADTSFTNRLOEMERISGAEDEIIDS
TKQDNVKKRKLQVNIQETQDSMRSNRIIGIESEDSPLGPNVFNKILIEENP
LKEEIPIDIOEAVRTPNRIDQKNTSRHIVTPNAONKERILKAVREKQVYKGRP
IRTPDEPSETWKARSWTDVITLREHKQPLLYPALSLINIDETKIFDKTFT
OYLSNPALORILNEFOENISKOVEARETOKSLKEFOEKIKOLKELMEIBAKK
EHMETTLDINQKQGVADTSFTNRLOEMERISGAEDEIIDSFTVDNKRKRL
VONTQDSKRSNRIIGIESEDSQKGVADTSFTNRLOEMERISGAEDEIIDS
PNRIDQKNTSRHIVTPNAONKERILKAVREKQVYKGRPILKAVREKQVYKGRP
QKQVSLINQKQVYKGRPILKAVREKQVYKGRPILKAVREKQVYKGRP
LAKTLHAPHTIIVGDEFTPLSMDSKQVILNSBVDLREVMQMDLDIRFTYP
KAKGYTPSAPHGTFPKIDHIIIGQKGLARVRIEIPCVLSHDKLQVNNKGM
PYTWKMNANLNDLVKEIKELIKNLEPENEDPTYSNIMDTMKALPNNKGM
ACRKOEAYVSSLTAHLKALKEKANTPRSRROETILRAEINOVETKRTIERINR
TKRWEKINKIDKPLARTGRHREVQINKIRNEGDIITVSEEOKILRSYKNLY
STKLENLOMDNPLDRQVSKNOBQINLNPIPEKEIAVYKGLPTKSGPDPDS
AEFYOTFIEDLPILSKLFKLETGALNSFEYEAITTIIPKHOTTKENRPSL
NINIDIKILANRIOEHKITLIHDQVPIPGMGWINRTIMIHAYIKKXON
HMITSLDAKAPDKQHPMIVLBERIGQBYLNTVRYKRPVANTILNBEKEAI
PLKSTROGCPSPYLPNIVLAVLBARIAQKEIKIQIGKEBVKISLPHMDITVLS
DPKSTSELKLIINFSKVAGYKINSKVAALYKKEBKEIRERTFTPIIDPNNIK
VLGVLTKQVDLYKNFETLKEIEEDRWMDLPCSGVIGINIVYMAILPRAHIF
NALPIKPIQFKELEDRITCKEILANNKERRAKAILNNRSGGILPIELKOYRVA
IKTAWYRNROIDQMNRIEDENMHTYGLIPDKGATIKOMKOSIPSKQPMWR
ATCRMOIDPCSPCTKLSKWKIKDLHITPDLKILEEGHLMHMGKQKPNTP
MAYALRSRIDKDLIKLOSFCAKQTVVTRKQPTDWEKIFNPTTERGLISKITVEL
KLDLRETNNPILKMGSELNKEFTLAEBCMAERHLKCCSTLIVRMOIKTYLFLUT
PCDWLRSKTOVADAGEDEVEKEHSIVGRNWKERPCPESTEBTIQOMWYITMEYVSA
IKNNEFMKFGVKMLENIIISLRTOSODIHAPLYLETILRKGHFLYMOGQGVVKA
RAPOSILNSAREBELPHFIKQGFVLRFHGIIQNQNHSLGLSSSWPARSGISTQORA
FLHSRGEAVGAEV"

ORIGIN

Query Match 8.7%; Score 749.2; DB 3; Length 6966;
Best Local Similarity 65.0%; Pred. No. 9.5e-156;
Matches 1274; Conservative 0; Mismatches 653; Indels 33; Gaps 10;

QY 1 TGTATGAAGCCATGTCACTTTAATACAAACCAGAAAGATAT--CAAAAAGAA 58
DB 4577 TCTACGAAGCCCACTTACTCTTATCTTAACCAACAAAGACAAACAAAGAGAG 4636
QY 59 ACTATAGACCACTACCTGATGATATATACATGACAGAAATCCCAACAAATCTAGCTA 118
DB 4637 ACTTCAGACCAATTTCCCTTAAGATATATGACGCAAAATATCAATTAATATCTGCA 4696
QY 119 ACCCAATCCACAGATATCAAGAGATATCCACCATTTGCAAGTGGTTTCATACCA 178
DB 4697 ACCGAATTCAGAGCAATCAAAACAATCATCCACATGATCAAGAGTTCATCCCG 4756
QY 179 GAGGTCA-GGATAGGTTAAATACACAGTCAATAATGTGATATCATCAATAAAGAA 237
DB 4757 GCATCAGGAGTGTATATATACGAAAAACCATGATCATCTTATATTAAGAAAC 4816
QY 238 TTAATAAACAATAATCATGATCATCTCAATAGATCTGAAAAAGCATTTGACAAATCT 297

DB 4817 TGAAGAAACAGAACCAATGATCATTTATTAAGATGCTGAGAAACATTTGACAAATTC 4876
QY 298 AACATTTCTTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGCATACCTTAATG 357
DB 4877 AACACCCCTTCATGATTAAGAGCTCGGAAGAAATGGAATTCAGGCCCATCTTAACA 4936
QY 358 TATATAAGCCATATATATGACGAGCCACAGCAACATTAATATGATGAGGGAAGATGGA 417
DB 4937 TGTATAAGCCATATATGACGAGCCACAGCAACATTAATATGATGAGGGAAGATGGA 4996
QY 418 AACATTTCTTCCTGAGAACTGAGCAAGCAAGATG-CTACTTTCACCATTTATTC 476
DB 4997 AAGCATCCCATTAATATAGGAGCTAGCAAGAGGTGCGCATCTCTCCATCTATTC 5056
QY 477 ACATAGTAGTGAAGTTTATGACCAAGCATATGACCAAGAAAGAAATCAAGGACACC 536
DB 5057 ATATAGTTCTTAAGATTTAGCCAGAGCATATGACCAAAAGAGATCAAGGGATAC 5116
QY 537 AATCAATTAAGAGAAAGTCAACTGCTGCTGATGATGATGATGATGATGATGATGATG 596
DB 5117 AGATGGAAGAAAGAGGCTAAATATATCACTATTGACATGACATGATGATGATGATGAT 5176
QY 597 AAAACCTTAAGACTCATATCAGAAAGCTCTAGAACTGATACATAATTCAGTAAAGTTT 656
DB 5177 GTGATCCCAAAAGTTCCACAGATGAGCTATAGCTGATTAACAACATTCAGCAAGTGG 5236
QY 657 CAGGATCAAACTAAATGTACCAAAATCAGTACGCTGATACCAACAGTACCAAG 716
DB 5237 CTGGGATTAATTAATTAATCACTAAATATGATGATGATGATGATGATGATGATG 5296
QY 717 CTGAGAAATTAATTAAGAACTCAAACTGTTTACATAGCTGTAA-----AAATCT 770
DB 5297 CGAGAGAGAAATTAAGGAAAGACACCTTATATATGACCAATTAATATTAAGTAC 5356
QY 771 TAAGATATTTCTTACCAAGAGAGTGAAGACCTTCAAGAGAAACTACAAACACAG- 829
DB 5357 TCGGTGTACTTAAACGAGCAAGTAAAGATCTGTACATTAAGAACTTCAAGACATGA 5416
QY 830 ---CTGACATCATAGATGACAAACAGATGAAACACATCCCATGCTCATGATGGTA 886
DB 5417 GGAAGAAATTAAGAAAGACCTCAGAAAGATGAAAGATCTCCCATGCTCATGATGGTA 5476
QY 887 GATCAATTTGTGAAGAAAGACATATGCGCAAGCAATCTCAAGATTCATGCAATTC 946
DB 5477 GGAATTAATTAAGAAAGACATATGCGCAATTTACCAAGCAATTCATGCAATTC 5536
QY 947 CCACCAATATCATCATCATCTTCAAGAACTGAGAAAGAAACATTTCTAAATTCATA 1006
DB 5537 CATCAAAATTCATCAATCAATCTTCAAGAAAGTTCAG-ACAGAAATTTGCAAAATTCATC 5595
QY 1007 TGGACCAACCAACCAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGCTTAGCAAAA 1066
DB 5596 TGGAAATAC-----AAAAACCAGATAGCTAAAGCTATCTCCCACAATA 5641
QY 1067 AGAACCAATCTGAGGATCATATTCATCTTCAAACTATACATCAAGGCTATATATCA 1126
DB 5642 AAAAGCTTCAGAGGGAATCACTATCTCGAAGCTCAAGAGATTAAGAGAAATAGGA 5701
QY 1127 CCAAAACATATGACATGATCAATTAATAAGGACATAGACCAATGAAAGAAAGAGAGA 1186
DB 5702 TAAAACTGATGATTTGTGACAGAAACAGACATAGACCAATGAAATGAAATTTGAAG 5761
QY 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCAACTGATTTTGAACAAGCAAAACAAAACA 1246
DB 5762 ACCCAAGAAATGAACCAACACCTATGTCATTTGATTTTGAACAAGAGCAAAACCA 5821
QY 1247 TAAAGTGGGAAAGACATTTCTAGTTAACAATAGTGTGAGATTTATGCAAGCAACAT 1306
DB 5822 TCCAAATGAGAAAGATGACATTTTCAGCAAAATGAGTGTGATTCATGAGAGGCAACAT 5881
QY 1307 GTGAAAGATGAACATGATCTGCTGCTCACTTAATACAAATGATGATCAAGATG 1366

Db 5882 GTAGAAGATGCAGATGCATGCATC-ATCACCCCTGTACAAAGCTTAAGTCCAGG 5940
 Qy 1367 ATCAAAGACTTAAATCTGAGACCTAAACCATTAATAATTTAGAGATATACAGAAA 1426
 Db 5941 ATCAAGAGACCTCCATCAACAGACAGACACTCAACTATATAGAGAAAATAGGGAG 6000
 Qy 1427 ATGCTTTCAGATTCATTCATTCGCAAGACTTCATGCGCAAGAACCCCAAGTAATGCA 1486
 Db 6001 CATCTGAGACAGATGGGCACTGGAATAATTTCTGTGAACAAACCAATGCTTATAGCT 6060
 Qy 1487 ACAAAAACAAAATAATATAGATGACTTAATTAATTAAGCTTTTGGCAGCAAAA 1546
 Db 6061 CTAAGATCAAGATTCGACAAATGGATCTCTATAAATCGCAAGCTTCTGTAGGCAAG 6120
 Qy 1547 ACAATCTTAGCAGACGACAAACCAACCCAGAGAGAAAATCTTCAAACTAA 1606
 Db 6121 GACACGGTGTATGAGCAAAACGGCAACCAACAGATGGGAAAAGATCTTTACCAATCCT 6180
 Qy 1607 GCATCTACTAGATCTAATATCCGGAATCCACAAAGAACTCAACAAATACGCAAGAG 1666
 Db 6181 ACAACAGAGAGAGGCTTATATCCAAATATATCAAAAGAACTCAAGAGTTAGACCGAG 6240
 Qy 1667 AAGCAACAATCCCATGAAAGAGTGGCTAAGACATGAATGACATTTCTCAAAAGAA 1726
 Db 6241 GAACCAATATACCTATTTAAATAATGGGTTCAAGCTTAACAAAGATTCACAGCTGAG 6300
 Qy 1727 GATATACAAATGGCCAAACAA--GGAATAATGCTTAACATCATATGATTAGGGA 1784
 Db 6301 GAATGCCGAATGGCTGAGAAACACCTAAAGAAATGTTCAATCTTATAGTAAAGGAA 6360
 Qy 1785 ATGTAATCAACATCTGTAATGCGATACCACTTACTCTGCAAGAAATGCTATATTTAA 1844
 Db 6361 ATGCAAAATCAAAACACCTGAGATTCACTCAACAC-ATGTATTTGGCTTAACATCAA 6419
 Qy 1845 AATCTAAATATATATGATGTTGGTGGTGTGTGTGTAAGAAACATTTACAGTGC 1904
 Db 6420 AACTCAGTACAGACATGCTGGCGAGATGTGAGAAAGAAACACCTCTCATTTGT 6479
 Qy 1905 TGGTGGAAATGTAACTTGGCGCAACCACTATGAGAAACAG 1944
 Db 6480 TGGCAGAACTGGAAGAACCCAGATGCCCTTCAACAGAG 6519

RESULT 10
 AY539894 3429 bp mRNA linear HTC 21-MAR-2004
 LOCUS AY539894 Rattus norvegicus LRRGT00143 mRNA, complete cds.
 DEFINITION AY539894 AY539894.1 GI:45478125
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3429)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.D., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Liver regeneration after PH
 Unpublished
 2 (bases 1 to 3429)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.D., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Direct Submission
 Submitted (03-FEB-2004) Henan Bioengineering Key Lab, Henan Normal
 University, NO.148 Jianshe Road, Xinxiang City, Henan 453002,
 P. R. China
 FEATURES
 source
 location/Qualifiers
 1..3429
 /organism="Rattus norvegicus"
 /mol_type="mRNA"

CDS
 /db_xref="taxon:10116"
 1..3429
 /note="Liver regeneration-related"
 /codon_start=1
 /product="LRRGT00143"
 /protein_id="AA566234.1"
 /db_xref="GI:45478126"
 /translation="MVQTRSEVYISQWESODERKNYTFNPLLIQTGPSPTVKLLN
 SVDTGMEKHKEVNDIYIHRSCIFKPKMKDSQVPLDISGSGIGDPASGP
 RAYVCRDITRONSFRDQICLSTLYMQRQYKLNQEIQLNANIIPKEIKTAVIK
 GLPTKSPGDDGSABEYOFIEDLPILEKFKHETDGLPSPFEATYTLIPKH
 KDPTKKNRPRPISLMNDIKLKIILANDEHKKTIHHNDQGFITGMOGMRTRT
 IWHYITNKLKQGNHNTISDAEKAPKTOHPNRIKLEIGIOPPLNTYKAIYSKP
 VANIKLNGEKLEAIPLKSGTRQCPSPYFNIVLEVALARIQOKIKGIQIEKVEY
 KISLFPADMTIVLSDPSSTRELLKLINNSKVAQYKINSKSVAPLYTEKQAKKEI
 RETTPEIIDNNIKYLGVTLLKQVLDLYNKNFPLREIEDLRLMDLPQSMGRIN
 IVMAALPKAIYFNPAIPKIPLOFPELRLRTCKFLNNKKPPIAAILNNKTSGG
 ITTPELKQYRAIVITKAWYVRDROIQDNRIEDPEMNPHTYGLIEDGAKTIQWK
 KDISFSKRCFPMNRSTCRMQIDPCLSPTCKLASKYIKDHIKEDTLKLEKGLKH
 EHWGTGKNPLNKPMAVALRSRIDKMDLILQSFCKAKDVTVRKQPTMEKIFVNP
 TTDRLGLISKLYKELKLDRETNPNPKKMSSELNKQPTAECEVAEHLKMCSTSLV
 REMQIKTTLEFHLTPVMMAKIKNSALFLIARKKEPRCPSTEEVIQOMWYIYMEYS
 AIKNNSMKFVGMKLEENITLSELQSODIHGHHLIARVSHGLMTVAGNVLWG
 GWESMTAPILGDHSGDRCGELOVNGAQYQDOVTPLOGATTLHPMPMPKPKPAK
 RIENHSRRATKQOEBBSVCKMEEDLIYNATLKLFTLLIRKGRYILASCIARIGINT
 FFFVFVRLHPONPCENAGDGPVAVLACSMQISRRYPSSSEPVSNQDSKNA
 OAAENRRETFMGDNDVAVDGSSTENKSNRQKELHIQIRASGHQROVTSWKAT"

ORIGIN
 Query Match 8.7%; Score 747.8; DB 3; Length 3429;
 Best Local Similarity 66.0%; Pred. No. 1.7e-155;
 Matches 1231; Conservative 0; Mismatches 602; Indels 32; Gaps 9;
 1 TGTATGAAGCAATGTCACTTAAATACCAAAACGAGAAAGATATA--CAAAAAGAA 58
 Db 617 TCTAGAGGCCAATTAATCTTATACCTTAACCAACAAAGACAAACAAAGAAAGAA 676
 Qy 59 ACTATAGACCGATACCACTGATGATATATATATGAGAAATCCCAACAAATTAAGCTTA 118
 Db 677 ACTTCAGACCAATTTCCCTTATGATATGACGCAAAATAATCTCAATTAATTTCTGCAA 736
 Qy 119 ACCCAATCCAAACGATATGAGAGATATCCACCATTTGCAAGTGGTTTCATACAG 178
 Db 737 ACCGAATTCAGAGACATCAAAACATATCCACATGATCAAGTAGCTTCATCCAG 796
 Qy 179 GGGTGA-GGATAGGTTAATACATACAAAGTCAATTAATGATATCATCATATAACAGAA 237
 Db 797 GCATGAGGAGATGTTTAATATAGGAAACCATCATATGATTCATTAATTAACAAAC 856
 Qy 238 TTAATAACAAAATACATGATATCTCAATGATGCTGAAAAAGCATTTGACAAATCT 297
 Db 857 TGAAGAAGCAAGAACCATATGATCTTCAATGATGCTGAAAAAGCATTTGACAAATCT 916
 Qy 298 AACTTTCTTATGATTAATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATG 357
 Db 917 AACACCCCTTCATGATTAAGATCTCTGAAAGAAATGGAATTCAGAGCCATATCCTTAACA 976
 Qy 358 TTAATAAGCATATATATGAGAGACCCACAGCAAAACATTTATGATGAGGAAAAAGTTGA 417
 Db 977 TAGTAAGGCAATATACAGAACCAAGTGTCTAATGATTAATTAATGAGAGAAACCTCG 1036
 Qy 418 AACATTTGCTCTGAGAACTGGAACAAAGACAGAGAG-CTACTTTCACCACTTATTTCA 476
 Db 1037 AAGCAATCCCATTAAGTTCAGAGGCTGAGCAAGGCTCCCACTCTCTCCCTACTTAATTTCA 1096
 Qy 477 AACTAGAGTGAAGTTTGAAGGAGCATATGAGAGAGAAAGAAATCAAGGGCAACC 536
 Db 1097 ATATAGTCTTGAAGTCTGAGCAAGCAATCAGACAAACAAAGAGATCAAGGGGATAC 1156
 Qy 537 AATCAATTAAGAGAGATCAAACTGCTGTTCACTGATGATATGATATACCTAG 596
 Db 1157 AGATCGAAAAGAGAGATCAAAATATCACTATTTGACAGATGATATGATATTTAA 1216

QY 597 AAAACCTTAAGACTCATCCAGAAAGCTCTAGAACTGATATACATAATTCAGTAAGTTT 656
 Db 1217 GTGATCCCAAGAGTTCCACAGAGACTCTAAGCTGATTAACCACTTCAGCAAGTTG 1276
 QY 657 CAGATACAACTTAATGTACAAATTCAGTACCTGCTATACACCAAGAGCCAG 716
 Db 1277 CTGGTATTAATTAATCACTAAATTAATCAAGTCCCTCTCTATACAAAAGAGAAACAG 1336
 QY 717 CTGAAATCAATCAAGAACTCAACCTTTTCACTAGCTGTA--AAATACT 770
 Db 1337 CCGAAGAAAGAAATTAAGGAAAGCAACCTTCTAATTAAGCCCAATTAATTAAGTACC 1396
 QY 771 TAAGATATTTCTTACCAAGAGGTGAAGGACCTCTACAGAGAACTACAAAACAG- 829
 Db 1397 TCGGTGCTTAACCAAGAGTAAGTAAGTGTATGATTAAGAACTTCAAGACCTGA 1456
 QY 830 ---CTGACATCATAGTGAACAACAAGTGAACAATCCCATGCTCATGGGTGA 886
 Db 1457 GGAAGAAATTAAGAGAACTCAGAAAGATGAAAGATCTCCATGCTCATGATGGCA 1516
 QY 887 GAATCAATTTGTGAAATTAAGACCATATTTGCCAAGAACTCAAGTTCAATGCAATTC 946
 Db 1517 GATTAATTAATTAAGAAATGACCATTTTACAAAGCAATCTAAGATTCATGCAATTC 1576
 QY 947 CCACCAAAATATCATCATCTCTTCAAGAACTAGAAAAAAACAATTTCTAAATTCATA 1006
 Db 1577 CCATCAAAATATCCAAATCCATCTTCAAGAGTTAG-ACGAACAAATTTGCAATTCATC 1635
 QY 1007 TGGAAACAACCAAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGACTTAAGCAAA 1066
 Db 1636 TGGAAATAC-----AAAAACCCAGATAGCTAAAGCATCTCAACACATA 1681
 QY 1067 AGAACAATCTGGAGGCTCATCATCCCATCTTCAAACTATCTCAAGGCTATATCA 1126
 Db 1682 AAGAGCTTCAGGGGAATCATCTATCCCTGAACTCAAGCAAGTATTAAGCAATGCTGA 1741
 QY 1127 CCAAAACATCATGCACTGACATTAACCTAGGACATAGACCAATGGAAGAGAGAGA 1186
 Db 1742 TAAAAAAGCTGATGTATTTGTACAGAGACAGACAGATGACCAATGGAATGAAATGAG 1801
 QY 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCCACTGATTTTGAACAAAGCAACAAACA 1246
 Db 1802 ACCCAAGAAATGAACCCACACACCTAATGCTCACTGATTTTGAACAAAGGCCAAACCA 1861
 QY 1247 TAAAGGGGGGAAAGCACTTCTAGTTAAAGAAATGAGTGAATTTGGCAAGCCCAT 1306
 Db 1862 TCCAAATGGAAGAAAGATGACATTTTCAAGAAATGAGTGTCTCACTGAGGCTCAACAT 1921
 QY 1307 GTGGAAGAAATGAATCTGATCCCTGTCTCTCACTTAATCAAAAATTTGATACAAATGG 1366
 Db 1922 GTAGAAAGATGAGATCGAT-CCATGCTTATCACCTGTACAAAGCTTAAGTCCAGTGG 1980
 QY 1367 ATCAAGACTTAATTAATCTGAGACCTTAACCAATTAATTTTGAAGATTAATCAAGAAA 1426
 Db 1981 ATCAAGAGACTCCACATCAAAACAGACACACTCAAACTAATAGAGAAAAAAGAGGAG 2040
 QY 1427 ATGCTTCTAGACATTCATTAAGCAAGATTCTGATGCGCAAGAACCCCAAGTAATGCA 1486
 Db 2041 CATCTGGAACACATGCGCACTGAGAAAAATTTCTTGAACAAACCAAAAGGCTTAATGCT 2100
 QY 1487 ACAAAAACAAAAATTAATAGATAGACTTAATTAATTAATTAATTAATTAATTAATTAAT 1546
 Db 2101 CTAAAGTCAAGATCGAATGAGATCTCAATTAATTAATTAATTAATTAATTAATTAATTA 2160
 QY 1547 ACAATCATTAAGAGAGCAACAGCAACCAAGAGTGAAGAGAAATTTTCACAAACTAA 1606
 Db 2161 GACACTGTGGTTAGCAAAACGCAACCAAGATTTGGAAGAAATCTTATCAAACTCT 2220
 QY 1607 GCAATCTGACTAAGACTAATTAATTCGGAATTCACAGAACTCAACAAATCAGAGAGAG 1666
 Db 2221 ACAACGAAATGAGAGCTTATTAATTCCAAAATTAATTAACAAAGCTCAAGAGTTAGACCG 2280
 QY 1667 AAGCAAAACAAATCCATGAAGAGTGGGCTAAGAGCATGAATGACAAATCTCAAAAGAA 1726

Db 2281 GAAACAAATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
 QY 1727 GATATCAATTAATGCGCAACAAACA--GGAAGAAATGCTTAACATCACTAATGATTAAGGAA 1784
 Db 2341 GAATGCGGATGCTGAGAGAAACCTTAAGAAATGTTCAACATCTTTAGTCAATAGGAA 2400
 QY 1785 ATGTAATCAACACTGTAATGCAATACCACTTACTCTGCAAGAAATGCTAATTAATTA 1844
 Db 2401 ATGCAATCAAAACAACCTGAGATTTTCACTCAACAGAGTGAATGCTTAAGTCAAA 2460
 QY 1845 AATTC 1849
 Db 2461 AACTC 2465

RESULT 11
 CR627265/c
 LOCUS
 DEFINITION
 CR627265
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 8575)
 Bloeker, H., Bloeker, M., Brandt, P., Mewes, H.W., Well, B., Amid, C.,
 Olsinger, A., Fodor, G., Han, M. and Wiemann, S.
 The German cDNA Consortium
 Direct Submission
 Submitted (13-JUL-2004) MIBS, Ingolstaedter Landstr.1, D-85764
 Neuenberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@kitz-heidelberg.de;
 sequenced by GSF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp686G01108) is available at the RZPD Deutsches
 Ressourcezentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686G01108
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.
 Location/Qualifiers
 1..8575
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686G01108"
 /rname_type="fetal kidney"
 /clone_lib="686 (synonym: hlec3). Vector pSport1_Sfi; host
 DH10B; sites SfiI + SfiIb"
 /dev_stage="fetal"
 /note="chimeric"

ORIGIN
 Query Match 8.6%; Score 744.8; DB 3; Length 8575;
 Best local similarity 70.3%; Pred. No. 9,4e-155;
 Matches 1221; Conservative 0; Mismatches 477; Indels 40; Gaps 15;
 QY 279 AAGCAATTTGACAAATCTTAACATTTCTTAAGATTAATTAACCTTGACGAAATGACATA 338
 Db 2922 AAGTTAATGAAGAAAGTTCACTTTTATTAATGAATTAATTAATCTCAAAAATCGGCTATA 2863
 QY 339 GAAAGACATACCTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 398
 Db 2862 GAAAGAACTACTCTCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2803
 QY 399 CTGAATGGGGAAGAGTTGAAGAACTTGTG-CCTGAGACTGGAACAGCAAGATG-CT 456
 Db 2802 CTGAATGGGGAAGAACTGAAGAGGCTCTCTTAAGATATGGAACATGATAGGATGCCCC 2743

QY 457 ACTTCACTCTTATTCACATAGTAGTGAAGTTTAAAGCAGACATCAGACAGA 516
 Db 2742 ACCCTCCACCTGTACTACATTTAGACCTGGAAGTCTTACTAGTAATCAACAGAGA 2683
 QY 517 GAAAGAATCAAGGCAACCAATCAATTAAGAGGAAGTCAATCTGCTCTTCACTGA 576
 Db 2682 GAAAGAAATTAAGGCACTCCAAACCGAAGGAAGACATCAAGTATCTTGTGTAGA 2623
 QY 577 TGATATGATGTATACCTAGAAAACCTTAAGACTCATCCAGAAAGCTCTAGAAGTAT 636
 Db 2622 TGATATGAATCTTATATTTGAAAAACCTTAAGACT--ACCAAAATCTGTCAATAGTAT 2565
 QY 637 ACATTAATTCATGAAGTTTACAGATACAACTAAATGTACACAAATACAGATAGTCT 696
 Db 2564 AAACAAATTCATGAAGTTTACAGATACAACTAAATGTACACAAATTCAGATAGTCT 2505
 QY 697 ATACACCAACAGTACACCAAGTGAAGTCAATCAAGAACTCAACACTTTACATAGC 756
 Db 2504 ATATGCCAAGGTACATCTGAAAAACAAATTTAAATTAACCCATTTATATAGTGC 2445
 QY 757 TGTAA--AAAAATCTTAAGATATCTTACCAAGAGGTGAAGACCTCTACAGGAA 814
 Db 2444 CACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2385
 QY 815 AACTACAAACACAGCTGAC-----TCATGATGACACAAACAGTGAAGACATCC 869
 Db 2384 AAATATTAACATGATGAAGAAATGAGAGACACAAACAACTGAGAGATATTC 2325
 QY 870 ATGTCATGATGGTGAATCAATATTTGAAGAAATGACATTTGCCAAAGAACTA 929
 Db 2324 ATGTTCATGATGAAGAAATCAATATTTGAAGAAATGACATTTGCCAAAGAACTA 2265
 QY 930 CAAGTTCATGATTCCTCCACCAAAATATCATCTATCTTCCACAGACTAGAAAAA 989
 Db 2264 CAGATTCATGATTCCTCCACCAAAATATCATCTATCTTCCACAGACTAG--GAAAA 2206
 QY 990 CAATTCATTAATTCATATGGAACACACAAAAAATTAATTAATTAATTAATTA 1049
 Db 2205 GAATTCATTAATTCATATGGAACACACAAAAAATTAATTAATTAATTAATTA 2160
 QY 1050 AGCAAGCTTAGCAAAAGAAACAAATCTGAGGATCACTTACCATCTTCAACTATA 1109
 Db 2159 AGCTATCTTAACAAAGAAACAAATCTGAGGAAACCACTACCTGCTTCAATATA 2100
 QY 1110 CTACAGGCTTATATCAACAAACATCATGCACTGACATTAATTAATTAATTAATTA 1169
 Db 2099 CTACAGGCTTATATCAACAAACATCATGCACTGACATTAATTAATTAATTAATTA 2040
 QY 1170 ATGGAAGAAAGAGAGAAATCCAAATTAAGCCAAATTAATTAATTAATTAATTA 1229
 Db 2039 ATGGAAGAAAGAGAGAAATCCAAATTAAGCCAAATTAATTAATTAATTAATTA 1980
 QY 1230 CAAAGCAACAAACAAATTAAGTGGGAAAGACATTTAGTAACTAAATGTGTGAGA 1289
 Db 1979 CAAAGTGGCAAGAAACATATCTGCGGAAAGACAGTCTTCAATTAATGTGTGAGA 1920
 QY 1290 TTATTTGCAAGCCATGTGGAAGATGAAGCTGATCCCTGTCTCTCACTTAATCAA 1349
 Db 1919 AATCTGATATCTATGTCAGAAAGATGAAGCTGAG--CCCTATATCTTCCACATATCAA 1861
 QY 1350 AATTTGATACAGATGATCAAGATCTTAATCTGAGACTTAACATTAATTAATTAATTA 1409
 Db 1860 AATTTGATACAGATGATCAAGATCTTAATCTGAGACTTAACATTAATTAATTAATTA 1802
 QY 1410 AAGTAACTACAGAAATTAATGCTTTAGACATTTAGGCAAGATTTCAATGGCCAGA 1469
 Db 1801 AAGTAACTACAGAAATTAATGCTTTAGACATTTAGGCAAGATTTCAATGGCTATA 1742
 QY 1470 ACCCAAAAGTAAAGCAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 1529
 Db 1741 CCCCAAA--TCACAAACAAACAAAGGCAAAATGACAAATGGAGTCAATTAATTAATTA 1683

QY 1530 GCTTTGGGAGCAAAAAAATCATCTAGACAGCAACAGAACCCAGAGTGAAGA 1589
 Db 1682 CTTTCTCAGCAGCAAGAAACAAATCAACAAAGTGAATTAACAAACCCGCAATGGAGA 1623
 QY 1590 AAATCTTCACAACTTAAGCATCTGACTAGGACTTAATATCCGAAATCCACAGAACTCA 1649
 Db 1622 AAATATCTGCAACTTACCCATTTGACAAAGGTTAATTAATTAATTAATTAATTAATTA 1563
 QY 1650 AACCAATCAGCA--AGAGAAAGCAACAAATCCATGAAAGATGGCTTAAGACATGAAT 1708
 Db 1562 AACCAACCTGTAGAAAAAATCTTAATTAATCAATCAAAAAATGGGCAAAAGATTGAAAC 1503
 QY 1709 AGCAATTCCTCAAGAAAGATATCAATATGGCCAAAC--AACGGAAGAAATGCTTACA 1766
 Db 1502 AGATATTTCTCAA--AAGACATACAAAGTGGCAACGCAATTAAGAAAGGTGCTCAATA 1446
 QY 1767 TCATTAATGATTAAGGAAATGTAATCAACATGTAATGCACTTAATCTCTGCA 1826
 Db 1445 TCATGATCATCAGAAAAATGCAAAATCAAACTTAATGACATCATGACCCAGTT 1386
 QY 1827 AGAATGCTCATTAATTAATAATTAATAATTAATTAATTAATTAATTAATTAATTA 1886
 Db 1385 AAATGGCTTACATCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1330
 QY 1887 GGAACACTTTTACCTCTGCTGGTGGGAATGTAACTTGGCAACCACTATGAAAAACGTG 1946
 Db 1329 GTAACTCTGTCACCTGTGGTGGGAATGTAACTTGTACAAACCACTATGCAATATGTT 1270
 QY 1947 TGGAAATTTCT--TAAGAACCTAAAGATGACATGATGATGATGATGATGATGATGAT 2003
 Db 1269 TGGAGTTCTCTCAAAAAAATTAATAATTAAGACTTCAATATGATCAGCAATCTCACTA 1212

RESULT 12
 AK038045
 LOCUS
 DEFINITION
 Mus musculus 16 days neonate thymus cDNA, RIKEN full-length insert sequence.
 ACCESSION
 AK038045
 VERSION
 AK038045.1 GI:26086252
 KEYWORDS
 HTC; CAP trapper;
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Harada, A., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohtsuka, E., Matsumoto, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, Y., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861

QY 1373 GACTTAATCTGAGACCTAAACCATATAAATTCTAGAGATTAACATGACGAAATGCTT 1432
 DB 1365 GAATCTCATATTAATAATGAGACAGTGAACCTTAGGAGGAAAGTGGGAAACCTT 1424
 QY 1433 CTAGACATTCATTCAGGCAAGACTCTCATGCGCAAGAACCCAAAGTAATGCAACAAA 1492
 DB 1425 GAGATTTTGGCACAGGGGAAAAATTCCTGATTAAGAACAGCAATGCTTGCTGTAGA 1484
 QY 1493 ACAAAATTAATAGATAGACTTAATTAATTAATAAAAGCTTTTGGCAGCAAAAACATC 1552
 DB 1485 TCGAGATTTGCAAAATGAGACTCATGAACTGCAAGACTCTGTAAGGCAAAAGACACC 1544
 QY 1553 ATTGACGAGCAAAACAGCAACCCAGAGTGAAGAAATCTTCAAACTAAACATCT 1612
 DB 1545 GTCAATTAAGACAAAAGGCCCAACAGATTTGGAAAGATCTTACCTATCTTAAGTCA 1604
 QY 1613 GACTAAGACTAATATTCGGAATCCACAGAGACTCAAAAGAAATCAGCAAGAAAGCA 1672
 DB 1605 GATGAGAGACTAATATTCATATATTAATTAAGAACTCAAGAGGTGAGCTCCAGAAATCT 1664
 QY 1673 ACAATCCCA-TGAAGAGTGGCTTAAGACATGATAGCAATTTCTCAAAAGAGATAT 1731
 DB 1665 AATTAACCCATTAATAAAATGGGCTCAGAGCTAAACAAAGATTTCTCACTGAGGATC 1724
 QY 1732 ACAATGGCCCAACAAAC--GGAATAATGCTTAACATCACTAATGATTAGGAAATGTA 1789
 DB 1725 CCGAATGGCTGAAGAAACCTGAAATAATGTTGACATTAATCAATCAGGAAATGCA 1784
 QY 1790 AATCAACACTGTAAATGGGATACCACTTACTCTGCAAGAAATGCTAATTAATTAATC 1849
 DB 1785 AATCAAAACAACCTTGATTCATCTCAACAGTCAAGATGGCTAAGATCAAAATTC 1844
 QY 1850 TAAATAATATAGATGTTGGTGGTCTGTGTGATTAAGAAACATTTTCACTGCTG 1909
 DB 1845 AGGTGACAGAGATGCTGGGAGATGTGAGAAAGGAAACATCTCTCCACTTTG 1904
 QY 1910 GGAATGTAACCTTGGCAACCACTATGSAACACGTTGCAAAATTTCTTAAGAACTAA 1969
 DB 1905 GGAATGCACTTGTATCAACCACTCTGAAATCACTGCTGGCAATTTCTCAAGAAATGGA 1964
 QY 1970 AGTAGATGACCATTTGATTCAGCAATCCC 1999
 DB 1965 CATATGTACTACCGAGGATCCGCAATACC 1994

 RESULT 13
 CA455045 889 bp mRNA linear EST 12-NOV-2002
 LOCUS AGENCOURT_10735784 MAFEL Homo sapiens CDNA IMAGE:6722551 5',
 DEFINITION mRNA sequence.
 CA455045
 ACCESSION CA455045.1 GI:24905369
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 889)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@rs-remail.nih.gov
 Tissue Procurement: Kristi A. Egland, Ira Pastan
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM14265 row: h column: 07
 High quality sequence stop: 743.

FEATURES
 ... source
 Location/Qualifiers
 1. 889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6722551"
 /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
 hTERT-HME1, LNCaP"
 /lab_host="EMD108"
 /clone_id="MAPCL"
 /note="Vector: PCMV-SPO6; Site 1: EcoRV; Site 2: Not I;
 Subtracted with brain, liver, lung, kidney and muscle.
 Directionally cloned. Priming method: oligo-dT. Average
 insert size: 1800 bp. Library amplification: 26,000 fold.
 Kristi A. Egland, James J. Vincent, Robert Strausberg,
 Bunkook Lee & Ira Pastan: Discovery of new breast
 cancer genes encoding membrane and secreted proteins.
 Manuscript submitted."

ORIGIN
 Query Match 8.2%; Score 708.2; DB 6; Length 889;
 Best Local Similarity 97.8%; Pred. No. 9.4e-147;
 Matches 727; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 7502 GGGTCATCAGACTGGGTTTCTGAGCATGATTTAACATCCAGTCTTGGGTACAGAAC 7561
 DB 147 GGGTCATCAGACTGGGTTTCTGAGCATGATTTAACATCCAGTCTTGGGTACAGAAC 206
 QY 7562 TGACCAATCAACGAGAGTGAAGAGACTCCTTCTCAAGCAGACCTGAGCTTCAAG 7621
 DB 207 TGACCAATCAACGAGAGTGAAGAGACTCCTTCTCAAGCAGACCTGAGCTTCAAG 266
 QY 7622 GGGTCATCAGACTGGGTTTCTGAGCATGATTTAACATCCAGTCTTGGGTACAGAAC 7681
 DB 267 GGGTCATCAGACTGGGTTTCTGAGCATGATTTAACATCCAGTCTTGGGTACAGAAC 326
 QY 7682 TGAGGCTCCGATGCGGAGGAGAGCGCTGTCTCATCTACATCTCAACCTGTCCGCGCG 7741
 DB 327 TGAGGCTCCGATGCGGAGGAGAGCGCTGTCTCATCTCAACCTGTCCGCGCG 386
 QY 7742 ACTTCCTCTCTTAAGGCGCACTATATGTTGCGGTTACGCTCATCAATATCCGCC 7801
 DB 387 ACTTCCTCTCTTAAGGCGCACTATATGTTGCGGTTACGCTCATCAATATCCGCC 446
 QY 7802 ATCCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCTTACTTATAGGCTTAAGCA 7861
 DB 447 ATCCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCTTACTTATAGGCTTAAGCA 506
 QY 7862 TGCTGAGCGCATCAGACCGAGGCGTCTGTTCATCTGTGCGCATCTGTGTAACACT 7921
 DB 507 TGCTGAGCGCATCAGACCGAGGCGTCTGTTCATCTGTGCGCATCTGTGTAACACT 566
 QY 7922 GCGCGCGCCCAATACCTGATCGGTCANGTGTGCTGCTGTGCGCCCTGCGCCGCG 7981
 DB 567 GCGCGCGCCCAATACCTGATCGGTCANGTGTGCTGCTGTGCGCCCTGCGCCGCG 626
 QY 7982 TGCGAGTATCTGAGTGTGATGTTCTGTGACTTCTGTTAAGTGTGATTTCTGTT 8041
 DB 627 TGCGAGTATCTGAGTGTGATGTTCTGTGACTTCTGTTAAGTGTGATTTCTGTT 686
 QY 8042 GGTGTGAAGCTGAGATTTTCAATCAATCGCGTGTGTTTATATGTGTGTTCTCT 8101
 DB 687 GGTGTGAAGCTGAGATTTTCAATCAATCGCGTGTGTTTATATGTGTGTTCTCT 746
 QY 8102 GTGGGTTCAGCTGTGCTCTGCTGAGATTTCTGTGTGATCCCGAAGATGCGCGCTGA 8161
 DB 747 GTGGGTTCAGCTGTGCTCTGCTGAGATTTCTGTGTGATCCCGAAGATGCGCGCTGA 806
 QY 8162 CAGAGCTGACGTGACCATCTCTCAAGTGTGTTCTTC-TCTCTGTGTGCGCTGCCC 8220
 DB 807 CAGAGCTGACGTGACCATCTCTCAAGTGTGTTCTTC-TCTCTGTGTGCGCTGCCC 866
 QY 8221 TTGGGATTCAGTGGGCTCTGTT 8243

Db 867 TTTTGCAATTCAGGGGCGCTT 889

RESULT 14
LOCUS BC028293 2776 bp mRNA 1linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:4704802, mRNA.
ACCESSION BC028293
VERSION BC028293.1 GI:22418059
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2776)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HWGC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
Kovats, C.R., Sneed, A.J., Martin, R.G., Wuzny, D.M., Nanavati,
A.N., Gibbs, R.A.

FEATURES
source
1..2776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4704802"
/issue_type="Breast", mammary adenocarcinoma."
/clone_lib="NIH MGC_87"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match 8.2%; Score 702.8; DB 3; Length 2776;
Best Local Similarity 72.9%; Pred. No. 1.9e-145;
Matches 1009; Conservative 0; Mismatches 347; Indels 28; Gaps 7;

628 AGAATGATACATAAATTGCAAGTTTCAGATACAAACTAAATGTACAAATCAGT 687
Db 864 AGAGTAGCTTCAAGCTTCAGCGAAGTCTCAGATATTAATCAATGTGCAAAAATCACA 923
Qy 688 AGCAGCTATACACCAAGAGTGAAGTGAAGTCAATCAAGACT--CAACACT 745
Db 924 AGCATCTTATACACCAATACAGACAGAGAGCCAAATCTAGAGTGAATCCCAATTCACA 983
Qy 746 TTACATATAGCTGTAAATAAATACTTAAGATATTTCTTACCCAGAGGTGAAGACTC 805
Db 984 ATTGCTTCAGAGATTAATATCTAGATCCATTCACAGGATGTGAAGAGACTC 1043
Qy 806 TACAGAGAACTACAAAACACAGCT---GACATCATAGATACACAAAAGTGGAAA 861
Db 1044 TTCAAGAGAACTACAAAACACAGCTCTCAAGGAATTAAGAGATACAAAACAATGGAG 1103

Qy 862 CACATCCATGCTCAGTAGAGGTAGATCAATATTTGTGAATAATGACCATATGGCAAAA 921
Db 1104 AACATTCATGCTCATGGGTAGAGAAATCAATTTCTGAAAATACATATGCCCCAG 1163
Qy 922 GCAATCTACAAATTGCAATGCAATTCACCAACAAATATCATCATCTTCTTACAGAACTA 981
Db 1164 GTATTTACAGATTCATATCCATCCCATCAAGTACCAATGACCTTCTTACAGAAA--TT 1222
Qy 982 GAAAAAACAATTTCTAAATTTATATGAGACAAACCAAAAAAATTTTTTTTTTCCCCG 1041
Db 1223 GAAAAAACAATTTCTAAATTTATATGAGAC-----CAAAAAAGCCAC 1268
Qy 1042 ATAGCCAAAGCAAGCTTAGCAAAAAAGAACAAATCTGAGGAGCATCAATCCACTTC 1101
Db 1269 ATTGCCAATCAATCTTAAGCCAAAAGACAAAGCTGAGAGATACAGCTACCTGACTTC 1328
Qy 1102 AAATCTACTACAAAGCTATTAATCAACAAACATCATGCGACTGCATPAAACTAGGAC 1161
Db 1329 AAATCTACTACAAAGCTACAGTAAACAAACAGCATGTGCTGTAACCAAAACAGAGAT 1388
Qy 1162 ATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAGCCAAATTAATATGCAACTG 1221
Db 1389 ATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAGCCAAATTAATATGCAACTG 1448
Qy 1222 ATTTTGAACAAGCAAAACAAACATTAAGTGGGAAAAGACATTCTAGTTAACAAATG 1281
Db 1449 ATCTTTGACAAACCTGACAAACAAACAAAGATGGGAAAGATCCCTATTTTAATTAATG 1508
Qy 1282 TGCTGATTAATTTGGCAAGCCACATGTGAGAGATGAATGAAATCTGATCTCTCACT 1341
Db 1509 TGCTGGGAAAATCTGGCTAGACATATGTAGAAAAGCTGAAATCTGATCTCTT--CTTACAC 1567
Qy 1342 TAATCAAAATTTGATACAAAGATGATCAAAAGCTTAATCTGAGACCTTAACCATTA 1401
Db 1568 TTAATCAAAATTTGATACAAAGATGATCAAAAGCTTAATCTGAGACCTTAACCATTA 1627
Qy 1402 AATTCTAGAAATTAATCATGAGAAAATGCTTTAGACATTCCTTAGGCAAAAGACTTCAT 1461
Db 1628 AGCCCTAGAAAGAAACCTAGGCAATACATTCAGACATAGGCAATGGCAAGAGACTTCAT 1687
Qy 1462 GGGCAAGAAACCAAAAGTAAATGCAACAAAACAAATTAATAGATGAGACTTAATTA 1521
Db 1688 GTCTAATAACCAAAAGCAATGAGCAAAACCAAAATTTGCAAAATGGGATCTTAATTA 1747
Qy 1522 ACTAATAAAGCTTTGGCGAGCAAAACAAATCTTGGAGAGAAACAGAACCAACCGA 1581
Db 1748 ACTCAAGAGCTTCTGACAGCAAAAGAAAGAACCAATCTGAGAGAAACAGCAACCTACGA 1807
Qy 1582 GTGAGAGAAATCTTCAACAACTTAAGCATCTGACTTAAGGACTTAATATCCGAAATCAGA 1641
Db 1808 ATGGAGAAATTTTGGCAATCTACTCANTCTGACAAAGGCTTAATATCAGAAATCTACA 1867
Qy 1642 GGAATCTAAACAAATCAGAGAGAGAAAGAAACAAATCCCATGAAGAGTGGGCTTAGGA 1701
Db 1868 TGAATCTAAACAAATTTTACAGAAAA---AAACAACCCATCAAAAAGTGGGCAAGGA 1923
Qy 1702 CATAATTAACAATTTCTCAAAAGAGATATACAATGGCCACAA--ACAGAAAAAATG 1759
Db 1924 TATGAACAAGACCTTCTCAAAAGAGAGCTTTATGCAAGCAAAAACATGATAAAAAATG 1983
Qy 1760 CTTAACATCAATATGATTAAGGAAATGTAATCAACCTGTAATGCGATACCACTTAC 1819
Db 1984 CTATATATCACTGGCCATCAGAGAAATGCAAAATCAAAACCAATAGAGATACATCTCAC 2043
Qy 1820 TCCTGCAAGATGCTATTAATTAATAATCTAATAAATAATAGATGTGGTGTCTGTGG 1879
Db 2044 ACCAGTTAAGATGCAATCATTAATAAAGTCAAGAAACAAACAGGTCTGGAGAGATGTAG 2103
Qy 1880 TGAATAAGAAACATTTTACATGCTGTGGGAAATGTAACTTGGGCAACCACTATGGA 1939
Db 2104 AGAAATAGAAACATTTTACATGTTGGTGGAGCTGTAACTATGTTCAACATTTGTGAC 2163
Qy 1940 AACAGTGTGAAATTTCTTAAGAACTAATAAGTAGATGACATTTGATCCAGCAATCCC 1999

```

Db      2164 GTGAGTGGGATTCCTCAGAGATCTTGAACATGAAATACCATTTGACCCAGTCATCCC 2223
Qy      2000 ATTA 2003
Db      2224 ATTA 2227

RESULT 15
CN839125      871 bp  mRNA  linear  EST 02-JUN-2004
LOCUS        AGNC009T15864180 NIH MGC 145 Homo sapiens cDNA clone
DEFINITION   IMAGE:30706609 3', mRNA sequence.
ACCESSION    CN839125
VERSION      CN839125.1 GI:47944780
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 871)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Daniela S. Gerhard, Ph.D.
              Office of Cancer Genomics
              National Cancer Institute / NIH
              Bldg. 31 Rm10A07 Bethesda, MD 20892
              Email: cgaabds-remail.nih.gov
              Tissue Procurement: GPCR Consortium
              cDNA Library Preparation: GPCR Consortium
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LMNL at:
              http://image.llnl.gov
              plate: IRB14 row: e column: 02
              High quality sequence stop: 701.
              Location/Qualifiers
                1..871
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30706609"
                /tissue_type="mixed"
                /lab_host="DH10B"
                /clone_lib="NIH MGC 145"
                /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
                varies by clone; ORFs were PCR-amplified and cloned into
                pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
                clone and include the following: 5'-EcoRV-XbaI/XhoI-3',
                5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional).
                For information about which gene each clone represents,
                please visit our anonymous ftp site at
                ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
                a Note: this is a NIH_MGC Library."

ORIGIN
Query Match      7.9%; Score 683.4; DB 7; Length 871;
Best Local Similarity 89.6%; Pred. No. 3.2e-141;
Matches 757; Conservative 0; Mismatches 86; Indels 2; Gaps 2;

```

```

Qy      7489 GTTTCGTGTTCCAGGGTCATCAGCTGAGGTTTCTGAGCATGATGATCAACCATCCAGTC 7548
Db      186 GTTTCGTGTTCCAGGGTCATCAGGACTAGGGTTTCTGAGCATGATGATCAACCATCCAGTC 245
Qy      7549 TTGGGTACAGAACTGACACCAATCAAGGACGTGAGAGACTCTTCTGTAACAACAGACC 7608
Db      246 TTGGACACAGAACTGACACCAATCAAGGACGTGAGAGACTCTTCTGTAACAACAGACC 305
Qy      7609 CTGAGCTTCAAGGGGCTGACGTGATGTTTCCCTTGTGCGCTGACAGAGAAAGCGGTT 7668
Db      306 TTGAGCTTCAAGGGGCTGACGTGATGTTTCCCTTGTGCGGCTGACAGAGAAAGCGATT 365
Qy      7669 GTGCTGTGGCTCTGAGGCTGCGGATGCGGAGAGACCGTGTCTCATCTACATCTTCAAC 7728
Db      366 GTGCTGTGGCTCTGAGGCTGCGGATGCGGAGAGACCGTGTCTCATCTACATCTTCAAC 425
Qy      7729 CTGCTGCGGCGGCACTTCTCTTCTTAGGGGCAATTAATGTTGCGCGTTACGGCTC 7788
Db      426 TTGGCGGAGAGACTTCTCTTCTTAGGGGCGGCTTAATTAATCTCTGTAGCTTC 485
Qy      7789 ATCAATATCGGCATCTCCATCTCCAAATCTCTAGTCTGTGATGACCTTCCCTACTTT 7848
Db      486 ATCAATATCGGCATCTCCATCTCCAAATCTCTAGTCTGTGATGATGTTTCTTACTTT 545
Qy      7849 ATAGGCTTAAGCATGCTGAGCGCCATGAGACCGAGCGCTGCTCATCTGTGAGCCC 7908
Db      546 GAGGCTGAGCTTCTGAGTGGCGGTGAGACCGAGCGCTGCTCATCTGTGAGCCC 605
Qy      7909 ATCTGTATCACTGCGCGCGCGCGCGAGATCTGTATCTGTATCTGTATCTGTCTCTG 7968
Db      606 ATCTGTATCACTGCGCGCGCGCGAGATCTGTATCTGTATCTGTATCTGTCTCTG 665
Qy      7969 GCCCTGCTGCTGCGGAGATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCTG 8028
Db      666 GCCCTGCTGCTGCGGAGATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCTG 725
Qy      8029 GCTGATTTCTGTGTGTGTGAACGTCAGATTTCATTAAATCGCGGTGTTTATTTT 8088
Db      726 GCTGATTTCTGTGTGTGTGAACGTCAGATTTCATTAAATCGCGGTGTTTATTTT 785
Qy      8089 TGTGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8147
Db      786 TGTGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 845
Qy      8148 GAAGA 8152
Db      846 GAAGA 850

```

Search completed: December 2, 2004, 18:05:18
 Job time : 22726 secs

This Page Blank (uspio)